PERIODICAL
GENERAL LIBRARY
UNIV. OF MICH.
1924

Journal of Educational Research

OFFICIAL ORGAN OF THE
EDUCATIONAL RESEARCH ASSOCIATION

B. R. BUCKINGHAM, EDITOR

ASSOCIATE EDITORS

E. J. ASHBAUGH — W. W. CHARTERS — GEORGE D. STRAYER
GUY M. WHIPPLE — S. A. COURTIS — LEWIS M. TERMAN
HARRY A. GREENE

VOL. X NOVEMBER, 1924 No. 4

CONTENTS

	Page
SOME PROBLEMS IN PUBLIC SCHOOL FINANCE.....Jesse B. Sears	257
ON THE FORM OF DISTRIBUTION OF INTELLECT IN THE NINTH GRADE.....E. L. Thorndike and E. O. Bregman	271
THE MERITS OF DIFFERENT METHODS OF SUBTRACTION.....J. T. Johnson	279
THE USE OF TEST DATA AS A BASIS FOR DRILL IN GRAMMAR.....H. Ward McGraw	291
VALUE OF GROUPING WORDS ACCORDING TO SIMILAR DIFFICULTIES IN SPELLING.....W. F. Tidyman and Edith Johnson	297
METHOD AND MEASUREMENT IN SPELLING.....Bertha M. Rogers	302
A LATIN COMPREHENSION TEST.....B. L. Ullman and T. J. Kirby	308
EDITORIALS	316
REVIEWS AND ABSTRACTS	321
NEWS ITEMS AND COMMUNICATIONS	329

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879.

Journal of Educational Research

Copyright 1924, by
Public School Publishing Co.

Official Organ of the Educational Research Association

Editor-in-Chief, B. R. Buckingham
Director of Bureau of Educational Research, Ohio State University, Columbus, Ohio

Associate Editors

E. J. ASHBAUGH, *Ohio State University, Columbus.*

W. W. CHARTERS, *University of Pittsburgh, Pittsburgh.*

S. A. COURTIS, *City Schools, Detroit, Mich.*

HARRY A. GREENE, *Ex-Officio, Secretary Educational Research Association, State University of Iowa, Iowa City.*

GEORGE D. STRAYER, *Teachers College, Columbia University, New York City.*

LEWIS M. TERMAN, *Leland Stanford Junior University, Stanford University, California.*

GUY M. WHIPPLE, *University of Michigan, Ann Arbor, Michigan.*

Contributing Editors

MURRAY DALMAN, *Director Department of References and Research, Indianapolis, Indiana.*

VIRGIL E. DICKSON, *Director, Bureau of Research and Guidance, Public Schools, Oakland and Berkeley, Cal.*

CHARLES FORDYCE, *Director of Bureau of Educational Measurements, University of Nebraska.*

WILLIAM S. GRAY, *Dean, School of Education, University of Chicago.*

M. E. HAGGERTY, *Dean of the College of Education, University of Minnesota.*

V. A. C. HENMON, *Director, School of Education, University of Wisconsin.*

ERNEST HORN, *University of Iowa.*

ARTHUR W. KALLOM, *Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.*

F. J. KELLY, *Dean Administration, University of Minnesota.*

ROBT. H. LANE, *Assistant Superintendent of Schools Los Angeles, Cal.*

WILLIAM A. McCALL, *Assistant Professor of Education, Teachers' College, Columbia University.*

GEORGE W. MELCHER, *Director, Bureau of Research and Efficiency, Public Schools, Kansas City, Mo.*

ARTHUR B. MOEHLMAN, *University of Michigan and Detroit Public Schools, Ann Arbor, Mich.*

EUGENE A. NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City.*

JOSEPH P. O'HERN, *Assistant Superintendent of Schools, Rochester, N. Y.*

W. J. OSBURN, *Director Educational Measurements, State Department of Public Instruction, Madison, Wis.*

J. B SEARS, *Stanford University, Stanford University, Cal.*

J. L. STENQUIST, *Director Bureau of Educational Research, Public Schools, Baltimore, Md.*

W. W. THEISEN, *Assistant Superintendent of Schools Milwaukee, Wis.*

C. W. WASHBURNE, *Superintendent of Schools, Winnetka, Ill.*

CLIFFORD WOODY, *Director Bureau of Tests and Measurements, University of Michigan.*

The JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public Schoo' Publishing Co. The subscription price is $4.00; $4.20 in Canada or Mexico; and $4.50 in other foreign countries.

The editorial office is at Columbus, Ohio. Address all communications on editorial matters to B. R. Buckingham, Director of Bureau of Educational Research, The Ohio State University, Columbus, Ohio.

Subscriptions and all communications about reprints or other business matters should be addressed to the Public School Publishing Company at Bloomington, Illinois.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume X NOVEMBER, 1924 Number 4

SOME PROBLEMS IN PUBLIC SCHOOL FINANCE[1]

JESSE B. SEARS
Leland Stanford University

I

The most outstanding problems in education today are problems of finance. Even though we are initiating important developments in education, developments that promise substantial changes in the curriculum and even in the general framework of the school system, yet, the attention of the public will not be diverted from the question of how we are to support this rapidly growing enterprise. Some critics affirm that we are attempting more than is socially desirable, educationally reasonable, or economically possible, while others suggest that we are wasteful, inefficient, and disorderly in what we are now doing however worth while our program may seem to be.

In the midst of these criticisms we find ourselves in a peculiar situation. The trend of school costs is still upward, yet the demand for free public schools is likewise on the increase. The public seems to buy what it wants and then to complain bitterly against oppressive taxes. We may say that all this is merely the unrest of a reconstruction period. We may point to the growing demand for more schooling, citing the fact that the school term has increased considerably during the past decade, and that attendance has grown at an enormous rate. We can point to the growing school budgets and to the expanding school debt, all as evidence that public education is safe.

[1] Read before the Third Annual Conference on Educational Research and Guidance at San Jose State Teachers College, May 23, 1924.

This, however, would be a wrong view. This is indeed an age of readjustment. All social institutions are being forced to join in this program of reconstruction and the public school cannot presume to be an exception. Whether we will or no the old objectives of education are rapidly giving way. The old curriculum and the old organization must find new objectives and be adjusted to them or, henceforth, the school will occupy a position in the social scheme of things secondary to that which it has so long held. It is not patience and watchful waiting therefore, but rapid, earnest, and constructive study that is needed. In this after-the-war mêlée, the free school will not tumble into its proper niche unaided. It must be put there by force of clear thinking and prompt action.

This will not merely call for a financial statement but for an explanation of the scheme of financiering by which the schools are maintained. The theories and principles which underlie our plans of school support are not well worked out, if we judge them by actual results obtained. All the children of all the people are not equally provided for. The burden of cost is not equally distributed. Education in the higher studies is not altogether free. We cannot pretend that our investments in school property are giving us a fair return in all cases. The turnover in teaching staff we know represents heavy loss. Our policy of going in debt is ridiculous in its waste. We know little about the cost of the subjects in our curriculum. Our methods of handling purchases, except in large systems, is inefficient and wasteful. We do not know what is a reasonable part of the budget for instructional service, for instructional supplies, or for any single item of overhead. We know that systems vary widely in all these items, but we do not know why they vary. We say it is due to differences in accounting and in school organization, but we do not know the merits of any particular claim. To be sure, we have made some studies of school costs, and we are not totally ignorant on any of these questions. Yet, we have almost nothing in the way of norms or standards in the entire field of school finance. We must admit that for the most part these are

unsolved problems and, further, that we have as yet no very large amount of scientific technique for working this field.

It is not the purpose here to present a long list of the problems of school finance nor to offer detailed analysis or solution for any one problem. It is rather to focus attention upon the situation as a whole, in order that we may more fully realize the need for investigation and research. This I shall try to do by sketching very briefly a few representative problems, from different parts of the field which would seem to offer suitable points of attack for such studies.

II

We need a clearer statement of our theory of school support. We have worked on the assumption that education is definitely a function of the state. The facts are that the state is in control, but that it is not the major source of support. In the United States, the state, the county, the township, and the local district all participate in the support and the control of the school. What should be the proper distribution of responsibility among these several political units? This is not merely a problem in finance —none of these problems is financial solely—but they are rather problems of administration and of social and political policy. The demand that equal educational opportunity be provided for all children cannot be realized by the present combinations of the various political units. Wealth is not distributed evenly throughout the different states nor in any direct ratio to school population or to school needs.

Neither do we have equality of cost burden in any state and seldom in any county. We have a good enough theory—tax wealth where it is to support schools where they are needed—but we do not live up to this theory if, indeed, we are not actually tending away from it. In 1910 the states supplied 26.9 percent of the cost of schools as compared with 25.7 percent in 1920.[2] This backward shift in the burden from state to county or to district means greater inequalities in cost burden. This shift has

[2] Newcomer, Mabel. *Financial Statistics of Public Education in the United States.* New York, Macmillan Company, 1923. 188 pp.

been due to the fact that the state contribution has not kept pace with the rapid increases in school expenditures during the past few years even though it has increased in the absolute.

We need look no further than to the maintenance cost of schools in the State of California for examples of these inequalities. In 1921-22 there was one county in this state in which the districts contributed nothing to the support of elementary or high schools. In a certain other county the districts contributed 48 percent of total cost. The lowest contribution made by the state in any county was 20 percent and the largest 48 percent of total cost. The county contributions ranged from 25 to 62 percent of total cost. Taking individual districts, we find that of the one-teacher schools of the state the individual districts contributed from nothing to 40 percent of the total cost. Taking the cities ranging in size from 5,000 to 10,000 population, the district furnished from nothing to 33 percent of the total school income.[3]

Thus, at its very foundation our plan of public education—the placement of responsibility for support and the equality of sacrifice for schools—needs to be re-examined. We have for decades been bringing our practice up in the direction of certain general ideals or theories. Now, because of the war and because the people are demanding more and more of the schools we are facing new and difficult conditions and there is some danger that the idea of free public education may receive a serious backset. If we wish to forestall such a possibility, two things at least are necessary. First, we must think these problems through. We must re-examine both our theories and our practice. Second, we must not only clarify our theory of public education but we must get this theory over to the people. It must become the accepted theory of the voter, the common citizen, and the taxpayer; as a people we must understand the foundation upon which the school rests.

Not only does it become obvious then that school administration must be well grounded in all aspects of public finance and

[3] See Sears, Jesse B., and Cubberley, E. P. *The Financing of Education in California.* New York, Macmillan Company, 1924.

political organization and management, but the general public must become better informed as to the nature of the foundation upon which their schools rest. Teachers must not only be better trained in the history and theory of the American free-school system, but they must teach in our schools something of the development of the ideals, and something of the present character and extent of the public school as one of the very most important strands in the story of our country, and as one of the most important—because constructive—factors in the civic development of our people.

There are other large problems which readily suggest themselves. We have invested large sums in plant, equipment, and maintenance of teacher-training institutions. Presumably our states have well-developed policies affecting the building up of a professional class of workers in education. We know that there is rapid elimination or turnover in this group. We know that the group represents wide diversity in extent and kinds of training. If we would make a few careful cost studies of turnover and of the returns in service we get for the various amounts invested in different degrees of training, it might help to rationalize our teacher-training program.

Again, we need some careful studies that will help to throw light upon our capital outlay investments. Almost nowhere do we find trustworthy figures showing the present value of the school plant. Year after year large sums are taken from the social income for building schoolhouses. We keep no track of the investment. What interest the money might have borne had it been left in private hands for reinvestment and the annual depreciation on the buildings are rarely or never figured. To some this kind of accounting for public outlays will seem to be unnecessary and even harmful. They say, "Let us forget the principal sum when it has been spent." My own judgment is that this is economically unsound, and further, sound business methods cannot be permanently harmful to education or to any other public service. Further, I am not so much concerned that every voter be asked to study such accounts. For those who are to furnish

leadership in the state, however, they are plainly necessary. The statesman cannot be satisfied with a partial view. He must have a full and clear perspective and this he cannot have if he overlooks this aspect of costs. A thorough study of how to keep track of such costs and how to figure the various kinds of depreciation on the various types of buildings and other school properties is, therefore, a present need.

The statutory control of public school debts and of outlay for buildings needs overhauling. In a neighboring state a district cannot erect a school building except by process of bonding. Many other states clearly encourage school debts by laws of a somewhat similar character. In like fashion we need better laws affecting the problems of budgeting school moneys. School budgets at present are largely a sham. Along with poor budgeting is poor control of accounts. Practically all these laws need to be studied as a step toward the development of higher standards along these lines.

III

Let us turn, however, from these broader problems of theory and practice in the large to some of the problems of financing an individual system of schools. It is here that most school people come into intimate touch with present-day criticisms, and it is here too that our own immediate responsibility usually begins. We should remember that every one of the problems we face in a small way in a single system is but a particular point at which these larger problems come to issue. For that reason we must think of many of these small problems merely as symptoms of disease that is deep-seated and far reaching and that, while we may get on very well in a given case by removing the symptom, the real cure involves much more. It reaches out into our statutes and beyond into our fundamental state and national policies and theories affecting education.

It may be said in general that wherever there is an administrative problem there is also a financial one, and that when administration has been reduced to a science we shall then be able

to proceed rapidly to a solution of most of the problems in school finance. Once we know what it is wise to do educationally, and also, how to administer the educational program decided upon, we can figure the cost and say whether it is or is not possible to pay for a given program, or which of several possible programs is desirable from the standpoint of cost.

At present many important educational questions are being settled on the basis of someone's best guess, not only as to educational merit and proper administrative procedure, but also as to the value and the cost of the product turned out. The fact that this idea of cost studies in education seems to frighten some who fear that our educational values and objectives may have to submit to economic standards need not alarm us. We should remember that it took educational tests nearly a decade to convince this same class of objectors that measuring educational processes and products is worth while.

The school-survey movement has developed the idea of efficiency studies in school administration and has contributed much in the way of analysis and technique for use in this connection. Yet an examination of city school reports and of current educational literature does not show that the idea of cost studies or of cost accounting has been carried very far. In this respect we must admit that school administration is somewhat behind administration in business, industry, and commerce, where every proposed change in policy and program is preceded by careful cost estimates and followed up by careful cost studies. Surely of two equal educational values we ought to purchase the one costing the least, and this we can do only if we know their cost as well as their educational value.

From the standpoint of the duties of school executives and school boards and also from the standpoint of the taxpayer and school patron, it would be possible to suggest a long list of studies that are needed. I shall suggest a few problems which will serve to typify the many that are necessary in any scientific administration of a school system.

First of all, the superintendent and board of education are constantly facing the question of raising more money for the various lines of development in school work, and the same is true generally of all public service. Not infrequently this means an increase in tax rate for schools and in time we have to decide whether the community is able to pay more. It is true that various studies have shown clearly that school support is rarely provided in terms of ability to pay but rather in terms of the social and educational ideals and aspirations of the people. Often the poorest district bears the highest tax and that by its own choice. If we had an accepted technique for determining the capacity of a city to pay for schools we would then be able to show in terms of standards just where a given city stands in the matter of ability to pay a higher school tax rate. Surely such knowledge would furnish the best means of changing the ideals of a community, when they were found to be below standard.

We can almost say that there is available at present in our survey literature a complete technique for such a study. We have not set up standards, however, and we cannot set up standards without further studies. The matter is not simple. Most of the comparative data on this point are only rough and many of the comparisons are misleading. One city receives 20 and another 30 percent of its school income from the state. The two do not contribute in the same proportion to the state and county school funds. Tax rates are high in one city because of investments in municipal enterprises such as gas or water service, whereas the other city has no such investments. There are differences in wealth, in earning power, in the ratio of school to total governmental needs, in bonded debt, and so on in almost endless complication. There is no doubt, however, that these differences can be ironed out if we will refine the technique we have. Once this is done then the development of norms is a relatively simple matter.

A second question and one which is of importance at the present time is that of when to bond and when not to bond. If a city's growth calls for one new building a year quite regularly,

we may well ask why not pay cash and save the interest and other costs connected with bonding. We have many city districts where the case against bonding is so clear that a child could understand it. Years ago, however, under other circumstances, bonding was probably begun in these cities and the tradition has grown up. It may be that there is something about the psychology of the situation that makes it easier to obtain school buildings by borrowing than by cash procedure. I am inclined to think that this is true but that it is not necessary for the schools to make progress that way. We know that public business is full of waste and often conducted by methods that are not sound. Surely our ideal is for a cleaner and higher type of public service. By introducing sound business methods in education we may hope gradually to accustom people to better things.

It is not only a question of when to bond, however, but also of how to bond. The merits and demerits of sinking fund methods for the liquidation of bonds is well known by a few students of finance and by some politicians but not by the general public. The same may be said of serial and long term bonds. We need more studies covering the actual experience of districts through long periods of time together with which are computations showing what would have happened by other methods. Out of such studies will come general principles and methods for studying any given case, methods which can be understood by the young superintendent and used in presenting his problems to his community.

A third problem that even a slight acquaintance with school-survey work will bring to one's attention is that of the cost of overhead in education. By overhead I mean all costs that are not directly connected with actual instruction. We know that among systems there are wide differences if we judge by the rather rough figures we find in financial reports. We need now to study the causes of these differences. Are they due to size of the system, to differences in organization, or to differences in efficiency of management? It is entirely conceivable that a careful analysis of overhead costs, followed by comparative studies of systems

operating under different types of organization, would throw much light upon the merits of the different types of organization. How large should an elementary or high school be before it pays to invest in a clerical assistant? What proportion of the salary budget should go for the service of school principals, for school supervisors, for research or staff service? In all these we find wide differences among school systems. We ought to study those differences from the standpoint of cost.

If we were asked to draw a floor plan that would be typical of the office of an elementary-school principal, or of a high-school principal, or of a superintendent, a registrar, a research bureau, or a business office, and then to list and describe suitable equipment and furnishings for each of these offices we should have great difficulty. We have no standards of any sort for any of these important features of the school plant and we do not know what their provision represents in overhead cost. We should quit using janitors' closets and scrap ends of corridors and we should not have too many magnificent suites for school offices. We should try to approximate standards which are as well suited to school management as are the scientifically organized and equipped offices of the large corporation in business or industry.

A fourth problem I might suggest is that of the cost of housing. Several questions readily come to mind in this connection. In a growing city how far in advance of present needs is it wise to build? Assume that we have a careful estimate of the school population for ten years ahead and have this data by individual communities. In putting up a new building for one of these communities how much ahead of present needs should we build? If this were expressed in terms of number of rooms, number of sittings, and number of dollars, we should have a clearer view of the problem. We must remember that unused rooms, seats, halls, toilets, roof, and furniture depreciate about as rapidly as if they were in use. True, we cannot always build with reference to number of children. We must often build with reference to the curriculum demands regardless of number of children. If so our question then becomes when this is necessary, What does

it cost over and above what it would cost if we built with regard to number of children? At some point this extra cost might be too great. When that cost seems high it will focus our attention upon the possibility of working out a building plan that can be added to unit by unit.

What is the best size for the rooms of a high school or elementary school or kindergarten? We have traditions and prejudices, but very little scientific basis for answering this question. Yet, it is a financial question of large importance. We need to know for each type of school and for each type of subject matter what is the best size of class. Second, we need to know how nearly possible it is to organize classes of maximum size. In the light of these two sets of facts we should determine the size and the number of sittings per room.

In a recent high-school survey it was found that, in a system including twelve buildings with sixty-seven classrooms, the smallest room contained ten sittings, the first quartile room contained twenty-eight; the median, twenty-eight; the second quartile, thirty; and the largest, sixty-nine sittings. The corresponding figures for the classes of the school were: lowest, three; first quartile, eighteen; median, twenty-three; second quartile, twenty-five; and largest, sixty-seven. From this, barring administrative difficulties—though we know that such difficulties cannot be barred—it appears that all classes could be comfortably housed. In reality it was not possible always to put the small class in the small room and the medium-sized class in a medium-sized room. At any rate, summing it up, it was found that, counting every room every day and every work hour the amount of service obtained was exactly 70 percent of the total service which the twelve buildings, with sixty-seven rooms, open eight hours per day represented. In other words, each dollar invested in plant was giving a return on a seventy-cent investment. When the matter was pursued still further, so that every seat was counted for every work hour, for every room in every building, then the dollar investment provided only a fifty-cent service. The question which this raises is, can we solve the administrative

problems that these facts suggest? Do we have to invest in a $100,000 building in order to get the housing service represented by $50,000?[4]

Staff turnover has been mentioned as a problem from the standpoint of the state in its program of teacher training. It is also a problem from the standpoint of the individual district. We do not know, but it is probably true that the smaller the school the larger the turnover in teaching force. The one-teacher school sometimes has a 200 percent turnover in a year. This is probably just as costly in education as it is in industry, and we ought to know both what it is doing to our schools and what it is costing us. Incidentally, in passing, it is not inconceivable that our tenure laws are making lack of turnover a problem in cost also.

We need to make cost studies of all our separate projects. We have schools of all sizes and all types. In large cities it is possible to have a school of almost any size we choose. We ought to be able to say that an elementary school of a given number of teachers and pupils is the most satisfactory from the standpoint of educational costs. If this size is in conflict with the size most desirable from the standpoint of educational results then we should be able to deal intelligently with the question of a compromise.

We need a clearer understanding of the problems of insurance, of repairs, of purchasing supplies, of managing supplies when purchased, of extra-curricular costs, of the cost of teachers' insurance and pensions, and of the cost of a scheme of teachers' sabbaticals. We need a larger knowledge of the process of budgeting and the problem of making inventories. We need detailed cost studies of each course—instructional and overhead separately—and of each separate project. Finally, too, we need much more efficient financial reporting than we now have, first, in order that school people may apply larger intelligence in the formulation of school policies and, second, in order that the public may have a clear and simple explanation of how their money has been used.

[4] Sears, J. B. "Some Data on the Cost of School Housing." *American School Board Journal.* Vol. 68, No. 6, June, 1924. pp. 42-44.

In order to make satisfactory headway in the solution of these problems and the many others they suggest there are two pieces of work that must be done. First of all, we must develop systems of school accounting which can be understood and which will make comparisons possible. This applies to our educational as well as to our financial statistics. In the past decade we have made much headway in the matter of standardizing child-accounting terminology and also in the matter of classifying school expenditures. This work must go forward. With the many statutory provisions affecting school accounting, with variety in sources of income, in methods of financing schools, in responsibility of state, county, and local school officials, we have built up a tangle of records that are most difficult to unravel and compare. This makes all our larger studies rough studies at best, and undoubtedly retards progress.

A second need is for a clearly defined terminology for use in cost accounting. The unit of cost is a plain necessity. Not one unit but many units, each carefully defined, and for some carefully specified function. "Per pupil," when used as a cost unit, must be defined and used only as defined. It cannot be wisely used for all purposes. We need to find out when and where it gives us a satisfactory measure of cost. One need not make a very extensive study of school costs in order to see how absurd it is to compare per pupil cost in a small school with per pupil cost in a large school. To be sure such a comparison has some value, but not the value that has so often been ascribed to it. So it is with "per teacher," or "per pupil class hour," or "per square yard of floor space," or "per cubic foot of room space," and other possible cost units. Only by extensive experimenting with these units are we going to be able to define their true mathematical functions.

IV

With these problems before us, some of them philosophical and others extremely practical in character, some of them state or national and others strictly local in their bearing, some of them large and others small, some of them intricate and very

difficult and others relatively easy of solution, we find ourselves face to face with practically a new field of study in education. These problems are but points of contact with this field. They do not suggest the orderly arrangement essential to a body of knowledge which we call a science. They are rather only promising prospects beneath which we believe are rich veins of ore.

It is true that financing, organizing, and administering are for education inseparable. A right financial solution must be a desirable if not always the very best educational solution: We must economize in education as well as in other lines of public service. Our aim should be to get full value for our money and not merely to save it. Not less expenditures but wiser expenditures should be our aim. Wiser expenditures will mean wiser policies and therefore better schools. If this kind of contribution can be made, then we shall in time have built up a body of well organized knowledge which we may be able to call the science of school finance or better perhaps, a branch of the science of public finance.

ON THE FORM OF DISTRIBUTION OF INTELLECT IN THE NINTH GRADE[1]

E. L. THORNDIKE AND E. O. BREGMAN
Teachers College, Columbia University

It has been generally assumed in the work that has been done heretofore in scaling tasks and determining units of difficulty, that the form of distribution in a grade or age group is that of the normal curve of error, a symmetrical bell-shaped curve, and the constants of this curve have been used in this work. Obviously, these constants are useful only in so far as the assumption is sound.

This investigation is an attempt to determine the true form for grade IX. We hope later to present data bearing on the form of distribution in other grades and in age groups.

The data used have been a series of frequency tables, representing the scores made by pupils in this grade in many cities of the United States on different tests of intelligence. The method in general has been to turn each table of frequencies into percentages to one decimal place, (or, in other words, to make the population of each group constant at a thousand) and to plot the percentage distribution from the mean as central tendency, and in units of one-tenth of the standard deviation of each distribution. The distributions were then superimposed, and the average frequency for each one-tenth sigma was determined. From this a final plot was made, which was thus a composite curve resulting from, and equally influenced by, each of the original curves. This final curve was then tested for goodness of fit to the theoretical normal curve

$$Y = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-x^2}{2\sigma^2}}$$

Our findings indicate that ability in the grade investigated is distributed in very close conformity to the normal curve.

A priori we may expect that a measurement in truly equal

[1] This investigation is one of a number made with the aid of a grant from the Carnegie Corporation.

units would reveal such a form in one of these late grades; the process of selecting pupils for any given late grade would tend to produce it. Ninth-graders, for instance, are selected because they are supposed to have attained ninth-grade ability; errors in judgment, however, and other chance errors would cause a certain number of pupils who really belong in grade VIII to be classified with the ninth grade, and a roughly equal number of pupils who really belong in grade X. Diminishing numbers would be included in each direction still more remote in ability from that set as the standard for the ninth grade. *Ergo* the bell-shaped curve.

The empirical evidence follows:

Figure 1 gives the composite curve which was obtained for the ninth grade. The broken line in the figure indicates the

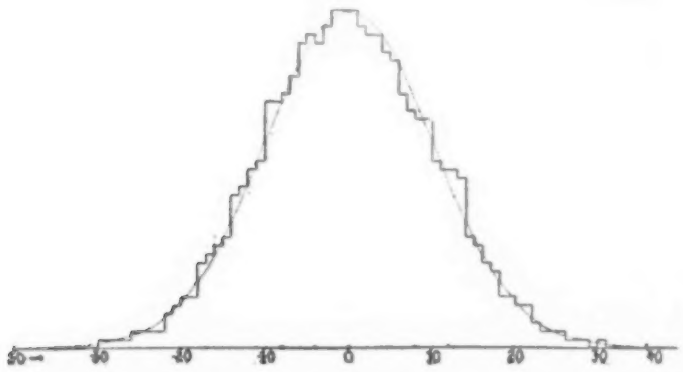

FIGURE 1. COMPOSITE CURVE FOR THE NINTH GRADE, BASED UPON NINE SINGLE CURVES. THE BROKEN LINE INDICATES THE THEORETICAL NORMAL CURVE

theoretical normal curve; the continuous line, the actual curve obtained from our data. The identity of the two is very striking. Testing the obtained curve for goodness of fit to the theoretical curve, we found P to equal unity.[2]

[2] Pearson, Karl, Editor. *Tables for Statisticians and Biometricians.* Cambridge, England, Cambridge University Press, 1914. pp. xxxi-xxxiii and Table XII (p. 26).

The single curves from which Figure 1 was derived by averaging, equal weight being attached to each, were plotted from frequency distributions for the following tests:

TEST	NO. OF CASES
I.E.R. Selective and Relational Thinking, Form A	3214[a]
I.E.R. Generalization and Organization, Form A	3214[a]
I.E.R. Total, Form A	3231[a]
Terman Group Test	1438[b]
Army Alpha	1721[c]
National A	494[b]
Haggerty	437[b]
Illinois Examination	380[b]
Otis Advanced Examination	232[b]

We do not show these individual curves for lack of space. Some of them were smooth, others showed notable valleys and humps. These irregularities may have been due to the numerical inadequacy of the samples, or to the inadequacy of the single measures of the individuals in a group, or to inequalities in the units of the tests, or to all of these causes, or any combination of them.

We are not at present concerned with the form of distribution of any of these curves, except to note that in general the larger the sample, the freer the curve is from valleys and humps, and the less it is skewed. The smoothness and absence of skew which we find for our composite would not diminish if the samples were larger.

Our concern is with the composite of all nine, wherein the inequalities in the units of one test will probably tend to be offset by inequalities in the opposite direction in some other test, since there are no known or conceivable forces whereby any two of

[a] These tests are fully described in the *Journal of Educational Research* for April, 1922. E. L. Thorndike, "Instruments for Measuring the Disciplinary Values of Studies."

[b] These data were compiled from material collected in connection with the Latin inquiry of the American Classical League, for use for another purpose. They represent many schools.

[c] These are the measurements collected in connection with our investigation of the psychology of algebra. *See:* "The Limits Set to Educational Achievement by Limited Intelligence," Table 4. M. V. Cobb, *Journal Educational Psychology*, November, 1922.

these tests, should, except by chance, have its units depart from equality in the same direction at the same locus of ability.

The tests are varied, their authors are many, and the purposes and grades for which they were devised differ. It seems most unreasonable to believe that the many men working at different times and constructing these tests with various ends in view, and not specifically with the testing of the ninth grade in mind, would so in concert have stretched and shrunk the units of their tests that a ninth-grade testing would give the "normal" curve, irrespective of the manner in which ability actually varies in grade nine.

There is one force other than the real distribution of ninth-grade pupils that might cause the symmetry and "normality" of the composite, namely, the inadequacy of the single measures of the individuals in the groups. If the individuals were really all alike, and the divergencies of these scores were *all* the result of the error of measurement, the form of distribution of the group would be that of the error of measurement. If, however, the normal curve does represent the true form in the ninth grade, a series of measurements of a group of ninth-graders from which the error had been in some measure eliminated, should be distributed in as close approximation to the normal curve as another series of measurements of the same individuals, in which the error was still present, or present in greater degree.

The sum of the scores on two tests, for instance, is a better measure of an individual's ability than the score on either one of the tests. By our hypothesis, a plot from a series of such combined scores would give as close a fit to the true curve as a plot from a series of single scores.

We have data of this character for a group of 1,659 ninth-grade students in a certain city. These students were tested in 1922 with Form A of the I.E.R. tests. A year later they were tested with Form B of the same tests, a similar and equivalent series. The tests of both forms have been grouped into performances of three kinds, ability with words, ability with numbers, and ability to deal with ideas of space. The scores on these special groups have been called the W, N, and S scores

respectively. We have then for each student 9 scores: the W score, the N score, and the S score for Form A and for Form B (the N score represents 29 minutes of work; the W, 32 minutes; and the S, 14 minutes); the sums of the W, N, and S scores for A and for B, each of which represents an hour and a quarter of work; and the total score for the two years, the sum of A and B, representing two-and-a-half hours of work for each student tested.

The A score or the B score should be a more adequate measure of an individual's ability than any single one of its components, the W, N, or S scores. Similarly, the sum of the A and B scores should give a better measure than either A or B alone. Then if the form of distribution of intellect in the ninth grade measured in truly equal units is that of the normal curve, a plot from the sum of the W, N, and S scores should give as close a fit to the normal curve as a plot from W, N, or S alone, and a plot of the sum of A and B should still maintain as close an approximation. Each successive summation serves to squeeze out more and more of the error of measurement, and of any spurious symmetry or "normality" due to it.

Figures 2 and 3 give the percentage distributions of the scores on the space tests of Form A, and the word tests of Form B

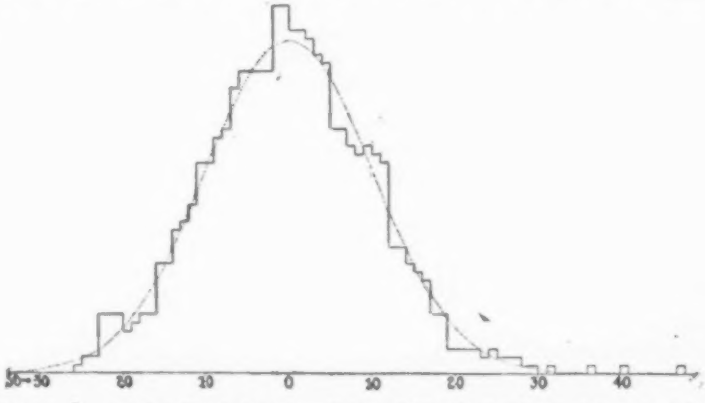

FIGURE 2. PERCENTAGE DISTRIBUTION OF SCORES IN THE I.E.R. SPACE TESTS, FORM A. $n = 1{,}659$. THE BROKEN LINE INDICATES THE THEORETICAL NORMAL CURVE

of the 1,659 pupils described above. We show the former as most typical of the distributions of the scores on the W, N, and S tests of Form A; the latter, as most typical of the three B distributions.

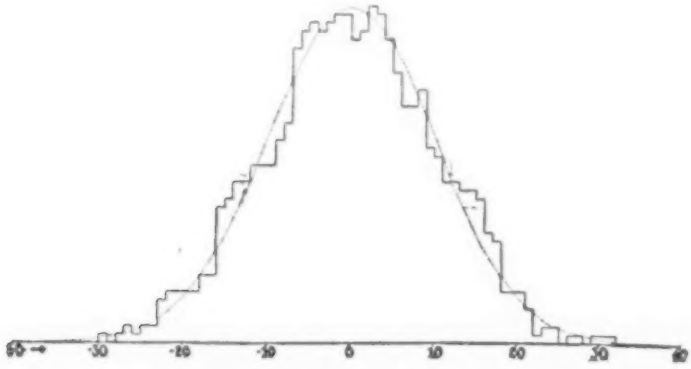

FIGURE 3. PERCENTAGE DISTRIBUTION OF SCORES IN THE I.E.R. WORD TESTS, FORM B. $n = 1{,}659$. THE BROKEN LINE INDICATES THE THEORETICAL NORMAL CURVE

The fit to the normal curve of each of the six distributions of the separate W, N, and S scores may be expressed as follows:

Scores	P (Goodness of Fit)	
	Form A	Form B
W	.712206	.958063
N	.999686	.989023
S	.809405	.878025

Figure 4 shows the percentage distribution of the scores on the whole of Form A, that is, the sum of the separate W, N, and S scores of Form A; and Figure 5 shows the percentage distributions of the scores on the whole of Form B.

The fit of Figure 4 is: $P = .999901$.
The fit of Figure 5 is: $P = .990295$.

Figure 6 gives the curve for the total score, the sum of the A and B scores. Its fit is expressed by: $P = .999550$. In Fig-

ures 4, 5, and 6, n equals 1,656, three cases having been lost in the course of the investigation.

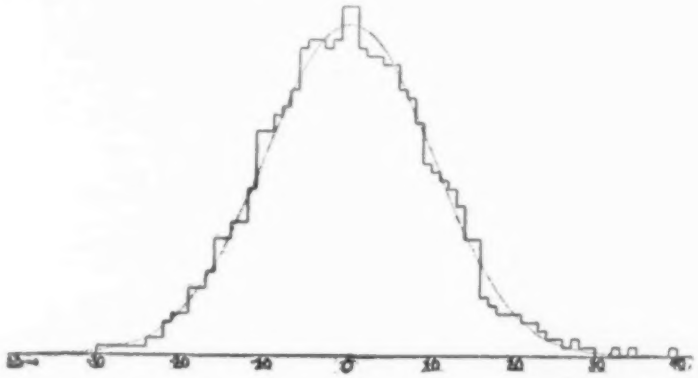

FIGURE 4. PERCENTAGE DISTRIBUTION OF SCORES IN THE I.E.R. TESTS, FORM A. THE BROKEN LINE INDICATES THE THEORETICAL NORMAL CURVE. $n = 1{,}656$

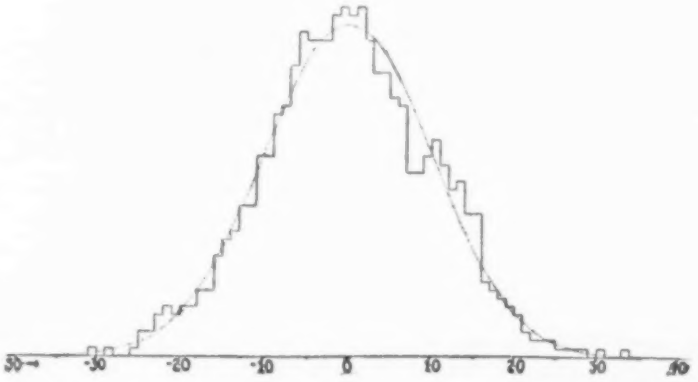

FIGURE 5. PERCENTAGE DISTRIBUTION OF THE SCORES IN THE I.E.R. TESTS, FORM B. $n = 1{,}656$. THE BROKEN LINE INDICATES THE THEORETICAL NORMAL CURVE

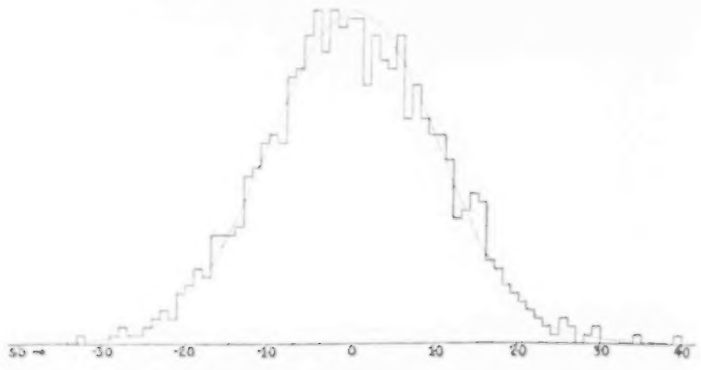

FIGURE 6. PERCENTAGE DISTRIBUTION OF THE SCORES ON THE I.E.R. TESTS, FORM A AND B COMBINED. $n = 1{,}656$. THE BROKEN LINE INDICATES THE THEORETICAL NORMAL CURVE

We are, therefore, led to the conclusion that intellect in the ninth grade, if measured in truly equal units, is distributed approximately in a surface of frequency bounded by

$$y = \frac{N}{\sigma\sqrt{2\pi}}\, e^{\frac{-x^2}{2\sigma^2}}$$

We shall, therefore, approximate closely to the truth by using this hypothesis in scaling the difficulty of intellectual tasks for a ninth-grade population.

THE MERITS OF DIFFERENT METHODS OF SUBTRACTION

J. T. JOHNSON
Chicago Normal College

In reading the literature on investigations in arithmetic, one is very apt to be confused by the processes in the various methods of subtraction. We have the additive method, the take-away method, the decomposition method, the Austrian method, the equal additions method, the complementary method, and others.

In a sincere effort to find out how many methods there really are, and the nature and merits of each, the writer made the following experiment during the year 1923-24 with 277 students in the Chicago Normal College. Lesson Card No. 33 from the Courtis Standard Practice Tests, Series B was used for each of 11 groups each of which had from 25 to 30 members. The cards, however, were not placed in the hands of the students. The examples were dictated, and care was taken to see that no one began to subtract before the signal to begin was given. Each test was administered by the writer. All the students seemed to be sincere in trying to find out which method was best. If there was any prejudice on the part of a student it was for her own method, and hence each one did her best.

After the 17 examples had been dictated, the scoring of time was explained to the students as follows: "After you have started at the signal which I will give, I want each one of you to keep on subtracting until you have finished all of the 17 examples. As soon as you have finished, look up at the board where I will be writing minutes and seconds like this. (Here the examiner illustrated by writing 1:05, 10, 15, ...55; 2:05, 10, ...55; etc.) If, for example, when you look up at the board after finishing the 17 examples, you find that I am writing 2 ..... 15, that will mean that you took 2 minutes and 15 seconds to complete your subtraction. Write down this time score in the upper right-hand corner of your paper."

This was explained again and questions were answered so that the method of procedure was thoroughly understood. The last instruction was that each one should subtract in her own way by the method she had always used.

After all the students had finished, the attention of each of them was called to the fourth example, which was:

$$\begin{array}{r} 137017 \\ 70719 \\ \hline 66298 \end{array}$$

Each student was then asked to analyze the method she used with this example as follows:

Either

9 from 17 are 8
1 from 10 is 9
etc.

or

9 from 17 are 8
2 from 11 are 9
etc.

or

9 and what are 17
1 and what are 10
etc.

or

9 and what are 17
2 and what are 11
etc.

or any other method used.

A careful explanation of each step was requested—a detailed showing of the exact procedure with every figure of the example. Example 4 was chosen as a good sample because it contained 9's, 1's, and 0's.

It was found that the time ranged all the way from 1 minute and 30 seconds to 7 minutes. Four distinct methods and some mixed ones were found.

First method.—By far the greater number (220 out of 277) used the following method:

9 from 17 are 8
1 from 10 is 9
7 from 9 are 2
0 from 6 is 6
7 from 13 are 6

This will be recognized as the good old method of our forefathers. It is pretty generally known as the decomposition method.

Second method.—Twenty-three of the 277, according to their own descriptions, used this method:

9 from 17 are 8
2 from 11 are 9
8 from 10 are 2
1 from 7 is 6
7 from 13 are 6

This is called the equal additions method by some investigators but not by all.

Third method.—Eight of the 277 used this method:

9 and what are 17
1 and what are 10
7 and what are 9
0 and what are 6
7 and what are 13

This is an addition method which is sometimes called the Austrian method, but there is also another method called the Austrian.

Fourth method.—Thirteen of 277 used this method:

9 and what are 17
2 and what are 11
8 and what are 10
1 and what are 7
7 and what are 13

Many investigators call this the Austrian method, and I think that it should be given this name. Some people, however, call it the equal additions method, since here also we increase the subtrahend.

It is evident that this terminology needs to be revised. A little study of these analyses will show that there are four distinct methods, of which two are subtractive and two additive. There are two variations of the subtractive method and two of the additive, since we may either decrease the minuend figure (borrow) or increase the subtrahend figure. One author has named these methods in the order of the above classification thus:

I. Take-away-borrowing
II. Take-away-carrying
III. Addition-borrowing
IV. Addition-carrying

In the results of this experiment they will be referred to simply as Methods I, II, III, and IV.

Thirteen of the 277 students used mixed methods, of which ten were combinations of I and III or of I and IV. The remain-

ing three were freak methods, of which two analyzed the fourth example thus: 9 from 18 is 9 and 1 less is 8. The rest of the procedure was as in Method I. The third user of a freak method gave this analysis: "9 + 8 = 17. I never can subtract any larger number from 1 or 0, and I have to go to a large digit in the top row and cross off, adding 10 and reducing by one to add a digit to the number to the right. 7 + 6 = 13."

Table I gives the details as to the time required to complete the test, the method used in performing the subtraction and the number of examples correctly worked. The table reads as follows (second line): Of the students who required from 1 minute 20 seconds to 1 minute 35 seconds[1] five used Method I, four getting 17 examples right and one 16; one used Method II and got 17 examples right; three used Method IV and got 17 examples right; and the average accuracy of the nine students who required this length of time was 16.9 examples. The reader will observe that a number in parentheses after a score shows the number of students who obtained that score.

The data given in Table I showed that the median time of those who used Method I was 35 seconds longer than that of those who used Method II, 20 seconds longer than that of those who used Method III, and 55 seconds longer than that of those who used Method IV. At the same time those who employed Method I were less accurate by 5.4 percent, 2.6 percent, and 6.3 percent respectively.

Although these differences may seem small, their significance is in fact considerable. When we recall that the average time was less than 3 minutes, an increase of 35 seconds and 55 seconds would mean a difference of about 25 percent and 45 percent. Moreover, since there were only 17 examples, the 5.4 percent and 6.3 percent difference in accuracy means that one less example was correct in Method I than in Method II or IV. In other words, we can say that the time of the students using Method I was one-fourth greater than that of those using Method II and nearly one-half greater than that of those using

[1] Really up to, but not including, 1 minute 40 seconds.

TABLE I.—THE TIME REQUIRED, METHOD USED, AND NUMBER OF EXAMPLES CORRECTLY COMPLETED

Time	Method I		Method II		Method III		Method IV		Mixed Methods		Average Accuracy
	No.	Scores	No.	Scores	No.	Scores	No.	Scores	No.	Scores	
1:00–1:15	2	17(2)									17
1:20–1:35	5	17(4) 16	1	17			3	17(3)			16.9
1:40–1:55	11	17(9) 16(2)	3	17(2) 15	1	16	2	17 14			16.5
2:00–2:15	30	17(10) 16(11) 15(5) 14(3) 10	4	17(3) 15			3	17 16(2)	1	17	15.9
2:20–2:35	26	17(11) 16(9) 15(2) 14(4)	5	17 16(3) 14	3	17(2) 13	1	17	1	16*	16
2:40–2:55	39	17(12) 16(17) 15(6) 14(4)	4	17(3) 16	1	15	1	5	2	16(2)	15.8
3:00–3:15	38	17(14) 16(11) 15(7) 14(5) 13	4	17(3) 14	1	16			2	17 17*	15.9
3:20–3:35	14	17(2) 16(6) 15(2) 14 13(2) 12	2	16 14	1	16	1	17			15.3
3:40–3:55	17	17(2) 16(9) 15(3) 13 12 11			1	16	2	17 16	1	14	15.3
4:00–4:15	13	17(6) 16(2) 15(4) 12							3	16 16* 11	15.6
4:20–4:35	9	16(3) 15(4) 12 9							2	16 15	14.5
4:40–4:55	6	17(2) 16 14 15 13									15.3
5:00–5:15	3	16(2) 12							1	14	14.5
5:20–5:35	3	17 14 12									14.3
5:40–5:55	1	14									14.0
6:00 and over	3	15 14(2)									14.3
Median Time		2:55		2:20		2:35		2:00		3:40	
Average Accuracy.		15.31		16.217		15.75		16.384		15.461	
Percent Accuracy.		90.0		95.4		92.6		96.3		90.9	

*—designates that a freak method was used.

Method IV, and that at the same time the users of Method I on the average did one more example wrong than did the users of Methods II or IV.

What conclusion can be drawn from these data? It seems to the writer that one conclusion suggested is that Method I, known as the decomposition method, though it is popular, is not the best method to use in subtraction. Among the others, the one to use will have to be decided by some other standard of merit, such as its teaching facility. It is obvious that II and IV are the best in speed and accuracy, and in both of them we increase the subtrahend. Other experiments will have to decide the relative ease of teaching these two methods.

The question of the validity of this experiment now arises. What about a control group? Were the groups comparable? The answer is that no control group was needed since the methods were not taught. As to the question of comparable groups, we may consider three factors: intelligence, native subtracting ability, and previous practice in subtraction. Of these factors the last is no doubt the most important, the second next, and the first the least important.

With these considerations in view, each student was asked to record her previous experience. It was found that none had taken a commercial course in high school. Indeed, they had all had practically the same course, the general course preparing for teaching. They had apparently had about the same training in the grades, six to eight years before this time. The grade experience of nine or ten years before was so far distant that it could hardly be operative at this time. Only two had had any recent subtraction experience; during the summer one worked as a clerk in a store, and the other as an accountant. It happened, however, that both of these students used Method I. In respect, therefore, to previous practice in subtraction, the students were fairly comparable.

It was impossible to measure the second factor, that of native subtracting ability.

The least related factor is that of general intelligence, for it is now pretty well agreed that there is little correlation between

adding or subtracting ability and general intelligence. But the question might be asked, "How do we know that the students using Methods II and IV were not the brightest or most able in arithmetic?" In an effort to answer this question the semester average in arithmetic for each student was obtained. It was found that the general average of those using Method I was 85.5, while that of those using Methods II, III, and IV was 86.9. This is too small a difference to be significant.

In anticipation of some other factor which might have affected the results, it was thought that perhaps the more intelligent students had selected the uncommon methods. Upon investigation it was found that only one had adopted a method different from the one in which she was trained. This was a bright girl (scholarship average, 95) who had been taught the decomposition method but had later changed to the Austrian (Method IV) for no apparent reason other than because she had seen someone else use it. Her time was 1 minute 30 seconds, with every example correct. On the other hand, scholarship average in arithmetic of *all* those who used a mixed method or a freak method was found to be 91.5. The meaning of all this seems to be that intelligence is not related to the choice of method in subtracting.

As the results from each section of students were tabulated, the median time and average accuracy for each section were also computed. It was found that in every group except one the results were better in both time and accuracy for Methods II and IV than for Method I. The summary only is given below:

TABLE II.—MEDIAN TIME AND AVERAGE ACCURACY OF THE ELEVEN GROUPS OF STUDENTS

Median of Group Medians or Averages	Method				
	I	II	III	IV	Mixed
Time	2:55	2:20	2:45	2:05	3:00
Accuracy					
No. of examples	15.9	16.3	16.0	16.9	16.0
Percent	93.3	95.5	94.1	99.2	94.1

From these results we see, as would be expected, the same relative ranking of the four methods as before. The most peculiar and interesting result is that among the four methods the fastest is the most accurate, the next fastest is the next in accuracy, and so on.

Aside from the main study another interesting feature was noticed incidentally. A glance at Table I will reveal that the 17's (indicating 100 percent accuracy) were bunched near the top of the table, the rapid-time end. A calculation was made of the average accuracy of results in each of the 20-second time intervals in the horizontal rows of the table (see column headed *Average Accuracy*). These were then grouped and the average was calculated for each group from 1:00 to 1:55, 2:00 to 2:55, 3:00 to 3:55, 4:00 to 4:55, and 5:00 and over. These averages were found to be respectively as follows: 16.8, 16.0, 15.4, 15.0, and 14.3. This reduced to a percentage basis is 98.7, 94.1, 90.8, 88.1, and 84.1.

This was a rather interesting result and could scarcely arise by chance. It corroborates the conclusion already established by Thorndike that the most rapid workers are also the most accurate.

Previous Investigations

In searching through the literature on this topic we find five or six articles which we may briefly summarize.

W. H. Winch[2].—Mr. Winch gave extensive training to thirty-eight pupils of the fifth and sixth grade in both methods—namely, the method of equal additions and the method of decomposition. He then examined the pupils by means of a series of tests. His results show only a slight difference in the amount of work accomplished by the two groups. The figures are 151.5 for the decomposition method and 152.8 for the equal additions method. He says, "It (the investigation) does show, however, that the poorer children profited by the equal additions method."

[2] Winch, W. H. "'Equal Additions' versus 'Decomposition' in Teaching Subtraction." *Journal of Experimental Pedagogy*, 5:207, 261, June and December, 1920.

Mead and Sears[3].—One of these classes had been taught the additive and the other take-away method. The teaching had continued for four months, 30 minutes a day being devoted to the work. The median of the take-away class was 4.5 above that of the additive class in the combinations and 0.7 above in the longer processes. It is evident that in this report the terms "additive" and "take-away" are used in the same sense as in this paper. Clearly there could be no method of decomposition and none of equal additions in simple subtraction combinations.

Joseph S. Taylor[4].—In 1919, about six years after Superintendent Maxwell of New York City had ordered that subtraction "by the addition process" should be taught in the schools, Superintendent Taylor raised four questions. These were:

a. What has been the effect of this method of subtraction?
b. Is it better than the other methods?
c. If so, why?
d. Do the children after six years of practice really use the method in their ordinary computations?

The results show that only 37.6 percent used the Austrian method, even after six years.

Superintendent Taylor answers only the last question. What was meant here by the Austrian method is shown by a quotation from the questionnaire that was to be answered by teachers of the fourth, fifth, and sixth grades:

"In doing this problem 61
—29

(a) How many say 9 from 11 = 2;
2 from 5 = 3?.....
(b) How many say 9 from 11 = 2;
3 from 6 = 3?.....
(c) How many say 9 and 2 are 11;
3 and 3 = 6?....."

Superintendent Taylor calls (a) the Italian method, (b) the second form of the Italian method, and (c) the Austrian method.

The results from the above questionnaire were:

[3] Mead, Cyrus D. and Sears, Isabel. "Additive Subtraction and Multiplicative Division Tested." *Journal of Educational Psychology*, 7:261, May, 1916.

[4] Taylor, Joseph S. "Subtraction by the Addition Process." *Elementary School Journal*, 20:203, November, 1919.

Method	Children using	
	Number	Percent
(a)	2,483	21.8
(b)	4,625	40.6
(c)	4,260	37.6

Since the teachers in New York were not all in favor of the Austrian method, it was perhaps not taught everywhere.

Therefore the results of this experiment are rather valueless for determining which is the best method.

It may be noted that method (b), which is our Method II, has the greatest percent of users, and that 62.4 percent of the pupils used a subtractive method, either (a) or (b), in spite of the order for an additive method. What an opportunity for finding the relative accuracy of the three methods, if the data were available!

W. W. Beatty[5].—In California where the addition method has been officially adopted Mr. Beatty found that the pupils still went back to borrowing. In his article, however, the distinction is not clear between the additive and the borrowing methods. The author does not take account of the fact that a "borrowing" method may be "additive" at the same time—like our Method III. Mr. Beatty gave a Courtis subtraction test to determine efficiency. It is to be regretted that he did not elucidate his methods as clearly as did Superintendent Taylor in his experiment. The results, however, as given by Mr. Beatty, are, in part, as follows:

Method	Number pupils using	Rate	Accuracy
Additive	52	8.2	81.7
Borrowing	123	9.2	79.3

These differences, besides being small, do not point in the same direction, since the better speed of one is offset by the better accuracy of the other.

[5] Beatty, W. W. "The Additive versus the Borrowing Method of Subtraction," *Elementary School Journal*, 21:198-20, November, 1920.

P. B. Ballard[6].—Mr. Ballard investigated 71 schools, of which 23 used the equal additions method and 47 used the decomposition method. He found that the former was superior. He says that with children 8 to 14 years old the equal additions method was superior by 10 to 40 percent. This experiment is valuable in that each school was tested in its own method, the one it was actually using. Where the method is changed factors carried over from the old method enter in.

W. W. McClellan[7].—Of 143 pupils, 11 to 12 years old, 63 were found to use the equal additions method and 80 the decomposition method. Mr. McClellan gave a series of tests for speed and accuracy, extending over 20 weeks. His results are:

Method	Rate (Examples in 10 minutes)	Percent of errors
Equal additions....	161	6.7
Decomposition.....	134	9.7

He concludes by saying, "In speed, accuracy, and adaptability the equal additions method is better, but in extended practice decomposition is better."

Summarizing these six investigations we find that three of them—those of Winch, Ballard and McClellan—gave results decidedly in favor of the equal additions method as compared with the decomposition method. One investigation, namely the one by Mead and Sears, gives an argument for subtractive rather than additive methods. The studies by Taylor and Beatty do not prove the merit of any method, because their conclusions are limited to the number of pupils using each method.

It appears that these researches agree rather closely with the results of this investigation in finding that the equal additions method (our Methods II and IV) is superior to the decomposition method (our Method I).

[6] Ballard, P. B. "Norms of Performance in the Fundamental Processes of Arithmetic." *Journal of Experimental Pedagogy*, 2:396 and 3:9, December, 1914, and March, 1915.

[7] McClellan, W. W. "An Experimental Study of the Different Methods of Subtraction." *Journal of Experimental Pedagogy*, 4:293, December, 1918.

Other experiments with larger groups of children of different ages and with more practice in subtraction under their own method, should be carried on so as to corroborate or disprove this inference.

Conclusions

1. There are four distinct methods of subtraction. They may be shown by either of the following classifications:

Classification 1	Classification 2
Subtractive	Borrowing in the minuend
I. Borrowing in the minuend	I. Subtractive
II. Increasing the subtrahend	III. Additive
Additive	Increasing the subtrahend
III. Borrowing in the minuend	II. Subtractive
IV. Increasing the subtrahend	IV. Additive

Methods that do not fall into one of the above classifications are either mixtures of two or more of these processes, or else they are freak methods. They should certainly not be taught.

2. The decomposition method is apparently not the best one to use in subtraction.

3. A method where the subtrahend is increased (by many called equal addition) appears to be better than one that involves borrowing in the minuend.

4. Further investigations will have to prove which variation of the equal addition method (II or IV) is the better.

THE USE OF TEST DATA AS A BASIS FOR DRILL IN GRAMMAR

H. WARD MCGRAW
Plattsburg High School

Whatever other purposes tests may serve, the classroom teacher is likely to look at them askance if she cannot see a way to apply the test data in the everyday work of teaching. Mention tests to her, and her reaction comes at once in the question "How can I use them?" If you expect to convert her to the cause, your answer must be definite and specific, and above all practical. She wants to be shown. Your answer must have an immediate bearing on some evident problem.

The most important of the problems which confront every teacher is that of adapting group instruction to individual needs. Every teacher knows that the progress of the group is dependent upon that of its most backward member, and that as far as a certain portion of the group is concerned, much of her work is unnecessary duplication.

The writer of a recent high-school English test, believing that "the young people of to-day must be drilled," has included in his book "hundreds of brief, disciplinary exercises in mechanics suitable for five- and ten-minute periods." The teacher who uses this text, or any other drill book, is confronted at the outset with the problem of determining what drill is most needed by the group, and what individuals most need the drill. She may, to be sure, go through the exercises seriatim, but it is submitted that such a method is wasteful and no more apt to be effective than a "shotgun" prescription. The proof of the pudding, however, is in the eating, as I am convinced by a testing experiment which is here described.

The Kirby Grammar Test was used. This test consists of a series of forty-three sentences each containing in parentheses a correct and an incorrect form, and a series of forty-three principles. The pupil chooses one of the forms in parentheses, and

in a designated space before the sentence writes the character which designates the principle involved.

I am quite aware that someone may legitimately ask if the whole troublesome problem of English grammar has been reduced to the simplicity of forty-three sentences and forty-three principles. Would that it had. As a matter of fact, these forty-three cases do cover most of the troublesome matters of functional grammar perhaps as well as any test can. But be that as it may, the significant point is that the test comes nearer to being a *standard test* of what the pupil ought to be able to do than anything the individual teacher is likely to make out.

The tabulating score sheet shows in what respects the class as a whole is weak, and also what individual pupils need particular attention. For example, a summary of the scores shows that in the sentence *Let James and (I) (me) do the work,* 21 out of 49 pupils chose the incorrect pronoun, and that 43 out of 49 did not choose the correct principle. In the sentence *(Its) (It's) a long road that has no turn,* only 3 out of 49 chose the incorrect form, and all of the 49 designated the correct principle. Obviously the class needs more drill on the construction involving the subject of an infinitive than it does on the distinction between the possessive pronoun and the construction *it's.*

In the conduct of the daily drill I proceeded upon the belief (1) that there is less necessity for drilling the group in matters where habits are fairly well fixed, and (2) that the individuals who ought always to be checked in the daily drill are those whose grades are below the median for the class. The tradition that every pupil must be called upon every day is in my opinion fatal to effective teaching. There are those whom you can safely neglect, just as there are those whom you cannot safely neglect. The problem is simply to find out who is who. Accordingly I put the burden of the classroom drill work upon the poorer pupils, using the better ones as a last resort in matters of particular difficulty.

The exact procedure I used may be described briefly. For each class I made out a table summarizing the number of errors

on each sentence and each principle. On a blank test sheet I recorded this information opposite each sentence. In addition to these data I entered numbers indicating the rule and exercise illustrating the principle, as given in the particular drill book used by the class. The first sentence in my record, with its data, is as follows:

7 P 5S 1. It is (I) (me). 34:XXIV

This sample reads as follows: Seven pupils failed to apply the correct principle, and five chose the incorrect form. Rule 34 in Wooley's *Handbook of Composition* states the principle, and Exercise XXIV gives the drill exercises.

The time given to drill was from five to ten minutes a day at the beginning of the period. The total time and effort, however, which each class put upon any one exercise was proportioned to the need as shown by the number of errors recorded before each sentence. Not infrequently I judged it necessary to devise exercises to supplement those in the drill book in cases where there was unusual difficulty.

The method which I have outlined is justified, I feel, by the data obtained by re-testing. It ought to be said here that after the test had been given to each class in October and the scores had been published, no reference was ever made to the test afterward. Any attempt to use the test as an incentive to better work or to the attainment of a better score "next time" can only result in making the pupils "test-wise" and destroying the objectivity of the test as a reliable measuring device.

Table I gives a summary of the scores made in English III by two sections of twenty-one pupils each. For example, the number of pupils having 42 or 43 principles and sentences correct is shown in the columns headed *Principles* and *Sentences* opposite "42-43" in the column headed *Number correct*.

This table shows that both of these groups (which it will be observed, were above the standard median for grade XI at the beginning of the year) made satisfactory progress. As a matter of fact, the median attainment at the end of the year was in advance of the standard for grade XII. The significant thing

in the table is the movement of the scores toward the top in the June columns. This indicates that the improvement in each class was general. The scores in the June *Principle* columns show

TABLE I. DISTRIBUTION OF SCORES MADE BY THE TWO SECTIONS

Number Correct	Section I				Section II			
	October 17		June 5		October 17		June 5	
	Principles	Sentences	Principles	Sentences	Principles	Sentences	Principles	Sentences
42–43	0	2	6	7	3	3	9	7
39–41	6	7	7	6	8	6	4	8
36–38	6	8	7	6	2	7	4	3
33–35	3	2	1	2	4	3	2	3
30–32	4	2	0	0	0	1	0	0
27–29	1	0	0	0	0	1	2	0
24–26	1	0	0	0	4	0	0	0
Totals	21	21	21	21	21	21	21	21
Medians	36.5	37.6	39.7	39.8	39.1	37.6	40.3	40.1

	Principles	Sentences
Standard median for Grade XI	33.8	37.3

TABLE II. REPEATED AND NEW ERRORS

	Total Error	Reduction		Errors Repeated		New Errors	
		Number	Per-cent	Number	Per-cent	Number	Per-cent
Section I							
October	273						
June	142	131	48	91	33	51	19
Section II							
October	274						
June	158	89	36	84	34	74	30

that both groups lack little if any of being entirely above the standard median.

An analysis of the errors revealed a condition which raised in my mind a question as to the efficacy of the method I used. The percent of *new* error which appeared on the June tests had seemed too large. The percents in Table II show the proportion

of repeated and new error in each section to the total initial error on the October tests.

It will be observed that in Section II the proportion of the new error on the test in June to the total initial error is 30 percent, whereas the reduction in the number of errors was only 36 percent.

During the year just past I tried with one section of English III the usual method of going through the drill exer-

TABLE III. RECORD OF A SECTION FOR TWO INTERVALS OF EIGHT MONTHS

Score	October, 1922		June, 1923		April, 1924	
	Principles	Sentences	Principles	Sentences	Principles	Sentences
42–43	2	0	7	5	11	6
39–41	7	10	7	9	3	9
36–38	3	5	5	8	5	8
33–35	2	6	2	3	5	2
30–32	4	2	3	0	1	0
27–29	2	2	1	0	0	0
24–26	4	0	0	0	0	0
9–11	1	0	0	0	0	0
Totals	25	25	25	25	25	25
Medians	34.5	37	39.4	39.3	40	39.6

Standard Medians	Principles	Sentences
Grade X	32.6	36.6
Grade XI	33.8	37.3

cises seriatim, spending on each exercise the time which in my judgment was necessary. The record of this section for two intervals of eight months is given in Table III. During the year 1922-23 the method used with Sections I and II was followed with this class, which was then a section in English II.

Table III shows that the progress of this group during the year 1922-23 was satisfactory. The April median indicates, however, that it did not during 1923-24 make the progress which might reasonably be expected in view of its previous record, and in light of the record made by Sections I and II in English III.

TABLE IV. REPEATED AND NEW ERRORS AFTER THREE TESTS, EIGHT MONTHS APART

	Total Error	Reduction		Errors Repeated		New Errors	
		Number	Percent	Number	Percent	Number	Percent
October	410	...	...	...	...	...	...
June	219	191	47	125	30	94	23
April	191	28	13	95	43	96	44

Table IV presents an analysis of the errors. This table speaks for itself. A comparison of the April test data with the June data for Section I demonstrates, I think, the superiority of the method used with Sections I and II.

VALUE OF GROUPING WORDS ACCORDING TO SIMILAR DIFFICULTIES IN SPELLING

W. F. Tidyman and Edith Johnson
State Teachers College, Farmville, Virginia

Some years ago a study of the value of grouping words according to similar difficulties in spelling was reported by Dr. Wagner.[1] Two hundred words which contained certain common spelling difficulties were chosen from the textbook used by the sixth grade. Two sixth-grade classes of approximately equal spelling ability, taught by teachers rated as equally proficient by the principal, were the subjects in this experiment. The words were taught in ten lessons; the hour and the length of the instructional period were the same for both classes. The method was also uniform except that in one class the words were grouped according to common difficulties and in the other class the words were taken up in chance order.

A preliminary test preceding instruction and final tests to compare improvement were given. The results show the superiority of the grouping method not only in improvement but also in the average number of mistakes made per pupil in the ten learning lessons. The rate of improvement made by the grouping class ranged from 68.2 percent in the preliminary test to 97.1 percent in the final test, while the non-grouping class rose from 73.3 percent to 93.1 percent. Using a proportional method of comparison and adopting "the unknown of the preliminary test" as the basis, Dr. Wagner calculated that the rate of improvement of the grouping class was 1.1 times as great as that of the non-grouping class. The average number of mistakes for the grouping class was 3.9 words per pupil, for the non-grouping class, 5.0 words. In his conclusions he states that grouping words into lessons according to spelling difficulty "secures better daily

[1] Wagner, C. A. *Experimental Study of Grouping by Similarity in the Teaching of Spelling.* Philadelphia, Pa. 1912. (University of Pennsylvania thesis.)

lessons, better final results, and greater steadiness or constancy of correct spelling."

The experiment, herein reported, was designed to test the value of grouping words according to similar difficulties in spelling and to substantiate if possible the results of the earlier experiment by Dr. Wagner. It was made in the fifth grade of the Training School at the State Normal School, Farmville, Virginia, under the direction of the author. The teaching was done by Miss Edith Johnson, who was the fifth-grade supervisor at the time. There were twenty-three pupils in the grade, and all of them took part in the experiment.

Eighty words, common to the vocabularies of fifth-grade pupils, and representing typical spelling difficulties were selected. These words were divided into two groups of approximately equal difficulty—the difficulty being determined by previous testing. The two groups of words are listed in Table I.

The Method

The general plan of procedure was to determine by preliminary tests the initial ability of the pupils to spell the words, to teach the words to the pupils by identical methods except in the matter of grouping, and by final tests to determine and compare the amount of improvement resulting from the two methods.

The preliminary tests of the two groups of words were given on successive days, the Non-Grouping List on February 17, and the Grouping List on February 18. It was found that out of a total of 920 spellings, 618 words in the Non-Grouping List, and 605 in the Grouping List were spelled correctly.

The forty words of the Non-Grouping List were presented to the class on the first four days of the week of February 24, the words of the Grouping List in the week of March 3. Ten words were studied each day.

The words of the Non-Grouping List were taken up in a chance order and studied according to the following plan:

1. The word was written on the board, pronounced, and used in a sentence or defined.

2. The word was written in syllables and pronounced slowly by syllables. The hard part of the word was picked out and underlined.
3. The pupils closed their eyes and visualized the word.
4. The word was spelled aloud by the class and by individual pupils.
5. The word was written several times.
6. After each word was studied in this way, all the words of the lesson were written several times.

The lesson consumed twenty minutes, and no work was done outside of class.

The words of the Grouping List were arranged in groups according to similar difficulties, and words with like peculiarities of form were studied in each lesson. The method of class-directed study was the same as for the Non-Grouping List except that in Item 2 the similarities of words having the same difficulties were pointed out. An equal amount of time was used, and the like conditions were maintained as when teaching the words of the Non-Grouping List.

Final tests were given on the Friday of the week in which the words were studied, and again twelve weeks later, May 23. The words were taken up in a *chance order in the final tests*, not in the order in which they were studied.

Results

Table I gives the results of the preliminary and final tests for both groups of words. The totals at the foot of the table show the number of correct spellings in each test. The differences between the totals on the preliminary and final tests give the gains. In the non-grouping method the gain in the Friday test is 207 words; May test, 178 words. The gains for the grouping method are: Friday test, 256 words and May test, 219 words. The percents of gain, found by dividing the number of words gained by the number of words missed on the preliminary test, are: non-grouping method, Friday test 68.5 percent, May test 58.9 percent; grouping method, Friday test 81.2 percent, May test 69.5 percent. The grouping method shows 13 percent more words learned on the Friday test than the non-grouping method, and 10 percent more words on the May test.

TABLE I. RESULTS OF THE TEST

Non-Grouping Method				Grouping Method			
	Preliminary Test	Friday Test[a]	May Test[b]		Preliminary Test	Friday Test[a]	May Test[b]
faint	19	22	23	smoking	14	19	20
fried	16	21	22	couldn't	11	21	22
fright	21	20	21	hitch	13	22	23
helpful	14	19	20	least	19	23	23
rough	18	20	21	threw	17	23	22
meadow	19	21	23	feast	20	23	21
pledge	14	20	20	shopping	13	21	19
table	23	23	23	wading	15	21	22
tunnel	16	20	22	isn't	12	19	19
strength	7	18	13	ditch	16	23	23
ourselves	13	21	21	surely	6	17	12
debt	16	22	22	meat	23	22	23
lain	11	19	18	mention	6	19	16
tied	21	18	22	screw	11	21	22
tight	21	22	20	hoping	17	21	20
trail	22	22	23	shouldn't	10	20	22
painful	14	20	22	skimming	10	23	19
tough	19	22	21	stitch	14	23	22
instead	12	20	17	sitting	17	22	19
piece	15	19	16	shade	20	23	23
acknowledge	3	20	18	wheat	22	23	23
stable	23	23	23	kitchen	19	21	23
level	11	20	21	shipping	15	21	20
length	8	19	15	brought	21	22	23
themselves	13	22	22	nicely	9	23	20
doubt	7	19	16	pitcher	15	22	22
dried	11	22	22	freeze	17	23	23
light	23	23	23	spade	23	23	21
straight	13	13	12	question	13	21	21
spoonful	12	22	22	rejoice	12	19	19
enough	18	22	21	spare	19	22	21
steady	13	21	14	sneeze	19	23	23
thief	18	22	22	shaking	22	21	21
truthful	9	17	22	sought	16	22	21
breath	19	20	14	ought	18	23	19
relieve	13	22	14	cheat	18	22	21
hedge	17	21	21	choice	9	22	15
thankful	14	22	22	losing	6	19	16
believe	22	23	21	drew	20	23	23
trouble	20	23	21	swimming	8	20	17
Number of Pupils	23	23	23	Number of Pupils	23	23	23
Total Right	618	825	796	Total Right	605	861	824
Gain over Preliminary Test		207	178	Gain over Preliminary Test		256	219
Percent Gain		68.5	58.9	Percent Gain		81.2	69.5

[a] Friday test immediately followed the learning exercise.
[b] May test was given eleven or twelve weeks after the Friday test.

Conclusions and Criticisms

In interpreting the results we must remember that all the factors excepting the grouping and comparison of words with similar difficulties were kept the same. These factors included: the pupils, the number, and difficulty of the words, the method of presentation excepting the one part under investigation, the amount of time devoted to learning, the order of words in testing, the time of the tests, and the personality of the teacher. The only qualification to this statement needed, is that the May test occurred twelve weeks after the conclusion of the study of the Non-Grouping List, and eleven weeks after the study of the Grouping List. The difference is probably insignificant.

With these facts in mind we may conclude that, within the scope of the experiment, grouping words of similar difficulty together in spelling is a significant factor, and that grouping words for study secures approximately ten percent better results than the study of words in a chance order. The superiority of the grouping method is shown not only in immediate learning, but also in the pupils' ability to spell the words after an interval of twelve weeks. The results of this experiment confirm in a remarkable manner the conclusions of Dr. Wagner in his *Experimental Study of Grouping by Similarity in the Teaching of Spelling*, for he concluded that the rate of improvement of the class using the grouping method was 1.1 times as great as the class not using the grouping method.

METHOD AND MEASUREMENT IN SPELLING

BERTHA M. ROGERS
Janesville, Wisconsin

The following study is the outgrowth of a spelling survey made in the elementary schools of a city system. The investigation proved that there was a wide range of spelling abilities within each grade and much overlapping between grades. As the methods of teaching the subject were, for the most part in accord with progressive practices, the root of the trouble seemed to lie in the grouping of the children. Because this was of common concern it was thought best before making any recommendations to conduct an experiment, in order to clarify the prevailing ideas in regard to the proper remedial methods and to demonstrate a workable procedure. An eighth grade was selected for this purpose, because, if a teacher of a grade which has the fullest curriculum and which includes pupils who differ widely in abilities, can meet individual spelling needs, then group and individual instruction is feasible in all the grades. This, Miss Lucy Whitmore, an eighth-grade teacher of the Lincoln School, demonstrated.

This study extended over a period of about sixteen weeks and included twenty-seven pupils who were in the first half of their eighth-school year. The investigation aimed to ascertain through measurement with Ayres' Measuring Scale for Ability in Spelling the spelling ability of each pupil at the beginning and at the end of the period, the extent of individual improvement within the period, and the results of instruction which aimed to meet individual needs.

RESULTS OF THE INITIAL TEST

The preliminary test consisted of the entire list of 1,000 words which comprise the Ayres' scale. Out of the 349 words found in the first twelve columns, Columns A to L, inclusive, the class misspelled no word more than five times. These words were put upon personal lists. The remainder of the scale, there-

fore, served as the real measure of the spelling ability of the group. Out of the 651 words found in the last fourteen columns, that is Columns M to Z, inclusive, the class misspelled 45 words nine or more times. For example, nine pupils misspelled *omit* in Column M, twelve pupils misspelled *except* in Column N, nine pupils misspelled *o'clock* in Column P, nine pupils misspelled *surprise* in Column R, and fourteen pupils misspelled *lose* in Column R.

The errors in spelling were due in the main to slovenly habits in the pronunciation of such words as *government, except,* and *accept;* to confusion in the parts of words pronounced alike but spelled differently, such as parts of *principal* and *principle;* to the omission of double letters in such words as *discussion, committee, necessary;* and to lack of recognition of the unphonetic elements in such words as *cordially, decision,* and *height.*

The preliminary test revealed that 11 percent of the class or three pupils were of fourth-grade ability; 11 percent of the class or three pupils were of fifth-grade ability; 22 percent of the class or six pupils were of sixth-grade ability; 41 percent of the class or eleven pupils were of seventh-grade ability; and 15 percent of the class or four pupils were of eighth-grade ability. The hatched portion of each line in Figure 1 records the standing of the pupil in the preliminary test.

Methods of Instruction

In order to meet the needs of each pupil, the teacher divided the class into three groups and designated them for convenience respectively as A, B, and C groups. Group A included pupils who were of eighth-grade spelling ability; Group B, pupils who were of seventh-grade ability; and Group C included all pupils who were below the seventh-grade in spelling ability. The forty-five words misspelled by nine or more pupils in the class were apportioned to groups B and C for class work in spelling; to Group C, the sixty words misspelled from six to eight times; and to Group A was apportioned only the words on each individual's spelling list. The individual lists of Group A included

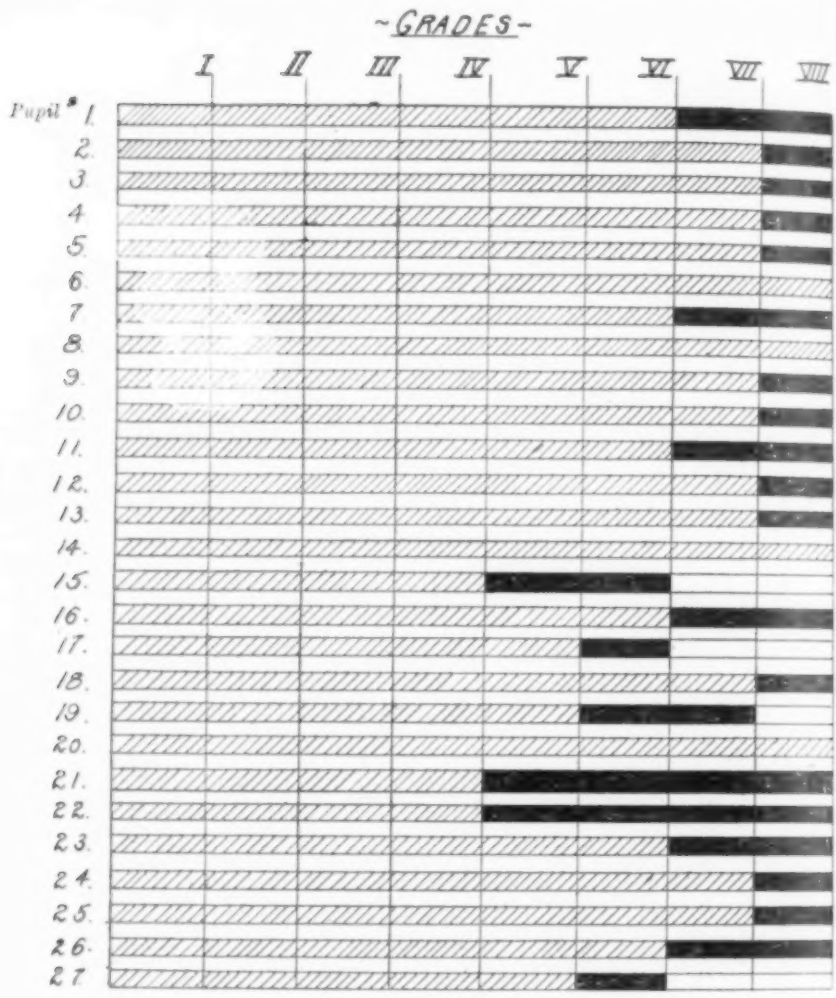

FIGURE 1. STANDINGS OF AN EIGHTH-GRADE CLASS IN SPELLING

from four to seventeen words. The personal lists of the individuals in Group B included those words which each had misspelled and which were not included in the list of forty-five words assigned to the group. These lists included from five to eight words. The personal lists of Group C included those words

which each had misspelled, and which were not included in the lists of forty-five and sixty words assigned to the group. These individual lists included from thirty-six to eighty-five words.

Thus Group A, was composed of four pupils who did no class work but learned the few words on their individual lists, while Group B, eleven pupils, received class instruction upon the forty-five words previously mentioned. This group worked to secure the advancement that would warrant their being classified as eighth- rather than seventh-grade pupils in spelling. And in addition, there was the third group of twelve pupils, who received class instruction upon the one hundred and five words previously mentioned. This group worked to attain successive grades beyond those indicated on the initial test with the hope of ultimately being classified as eighth-grade pupils in spelling. As a further incentive to effort each pupil kept a record of the results of his regular tests in the form of a graph.

Once a week the B and C groups together had a lesson on some of the forty-five words which in the initial test the class had misspelled. Group C had lessons twice a week also on some of the sixty words which in the initial test the class had misspelled. Thus each week the C Group had three lessons and the B Group, one lesson. Each consisted of four new words and required a twenty-minute period. The test on daily lessons embraced the four new words and the sixteen words presented in the four preceding lessons.

This plan made it possible to cover the entire class list of one hundred and five words in about nine weeks. This limit, however, did not include the time devoted to testing.

Three times the teacher tested the entire class on the 651 words in the last fourteen columns of the scale. Furthermore, when pupils failed to reach an eighth-grade standing on a given column, they continued to study the troublesome words in the particular column until in three tests they attained the eighth-grade norm on that column. This necessitated the testing of some children six times on one or more columns.

During the sixteen weeks the four pupils in Group A were by no means idle. They helped their teacher and classmates, for when a pupil in Group B or C wished to be tested on his personal lists or on a given column of the scale, a pupil in Group A who had assumed responsibility for that pupil's advancement, dictated the list, and then corrected and graded it. These pupils also assisted the teacher in recording the results of tests and in supervising independent study.

As an aid in this kind of study the teacher provided each child with directions suggested by Dr. Horn in an article entitled "Principles and Methods in Teaching Spelling as Derived from Scientific Investigation."[1] Guided by the suggestions contained in the nine rules of Dr. Horn, the pupils gradually acquired skill in memorizing for they devoted forty minutes a week in school, and in addition, voluntarily gave time outside of school to the study of their individual lists.

Results of Final Test

At the close of the sixteen-weeks period the teacher gave to the entire class a test of fifty words which for Groups B and C had proved especially difficult. For this test she selected all the words from the third test on Column M to Z inclusive which the class had misspelled three or more times and of those misspelled two times, enough to complete the list of fifty words.

The first quarter of the list fell in the middle, Column S, which sets for grade eight a norm of 92 percent, for grade seven a norm of 84 percent, for grade six a norm of 73 percent. These the teacher used as standards for the final test.

Out of the twenty-seven pupils who took the final test, twenty-three or 85 percent showed eighth-grade ability; one pupil or 4 percent, seventh-grade ability; and three pupils or 11 percent, sixth-grade ability.

Comparisons of the initial ability and the final score of each pupil are shown in Figure 1. The hatched portion of each line

[1] *Eighteenth Yearbook of the National Society for the Study of Education.* Part II. Bloomington, Illinois, Public School Publishing Company. 1919. p. 72.

records the grade of ability revealed by the pupil in the original test, while the solid portion shows his score on the final test. From this graph it appears that all of the pupils whose original test showed them to be of seventh-grade spelling ranking advanced during the trial period to the eighth-grade standard, while all the pupils having an original rating of sixth-grade rank, and two pupils of fourth-grade advanced to eighth-grade standing. The advancement of these nineteen pupils ranged from one to four school grades; one pupil of fifth-grade standing advanced to seventh, one of fourth-grade to sixth, and two of fifth-grade to sixth. Twenty-three or 85 percent attained or exceeded the norm for the eighth grade.

The study gave satisfaction to the teacher because it established for herself and for her pupils specific and attainable goals. The recognition of these goals aroused in the pupils a sense of responsibility, while it awakened appreciation of the extent of individual variation within the class and a desire for methods which meet special needs. The study, also, provided a correct measure of individual accomplishment and thus stimulated both teacher and pupils to further effort in the direction of worthwhile study. Further, the study was gratifying for it yielded an economical and a workable plan, the following of which raised the standing of the whole group as well as of the lowest half and in addition narrowed the range of the spelling abilities within the class.

A LATIN COMPREHENSION TEST

B. L. Ullman and T. J. Kirby
University of Iowa

The development of a new type of Latin test perhaps deserves a special record, particularly as it has seemed to find favor with the teachers and specialists in the fields both of Latin and of measurements. The construction of the test was undertaken by the authors for the American Classical League with the support of the General Education Board.[1]

This Latin Comprehension Test[2] had its origin in the authors' belief that one of the objects of the study of any language, foreign or vernacular, is to comprehend clearly and rapidly the meaning of the printed page. The number of silent-reading tests now available, the emphasis on silent-reading methods, and the development of reading tests in special subjects such as history and general science indicate the general acceptance of this belief for English.

In planning the test it was at first thought best to construct four parts, one for each year of the high-school Latin course. After careful consideration it was concluded, however, that a single test, flexible enough for use by all, was more desirable since it readily made possible a comparison between the results of the teaching in the different years. Several types were considered: (1) a test consisting of numerous sentences, (2) a test made up of a single long passage, or (3) one composed of a number of short paragraphs. The first was rejected because it failed to test the pupils' ability in reading connected discourse. The second was impossible in a test which was to be given to pupils who had studied Latin for varying periods of time, because it would necessitate the preparation of special material instead of the use of existing tests. The result would be at variance with the classroom conditions. The third type, a compro-

[1] Valuable assistance was obtained from a member of the Investigating Committee, Professor W. L. Carr.

[2] The test is published by the Extension Division, University of Iowa, Iowa City, Iowa.

mise between the first two, proved the most acceptable. The closest parallel among the tests of English comprehension is that of the Thorndike-McCall.

In the construction of a foreign language comprehension test, however, conditions are not the same as in the formation of an English comprehension test. No test in a foreign language can be as purely a comprehension test as one in the vernacular. For example, the maker of an English comprehension test for English-speaking children may take for granted a considerable vocabulary and some understanding of grammatical forms and constructions on the part of even a young pupil; but the vocabulary and grammatical difficulties of Latin make it necessary, when constructing a Latin comprehension test, to select materials which are within the student's range in these respects, as well as in thought content.

Preparation of the Test Items

A short paragraph of very simple Latin was composed and given in mimeographed form to several groups of high-school pupils. Ten questions were asked on this paragraph. The time taken by the pupils was carefully recorded. The same pupils then translated the passage and the translations, while not graded, were studied for their bearing on comprehension. This passage, translated into English, was then given to a group of pupils in the seventh grade to determine whether the questions were too simple or too difficult from the standpoint of comprehension without the added difficulty of the foreign language. As a result of these practice testings, the paragraph was cut down and became the "practice paragraph" used in the final form of the test. The number of questions was reduced from ten to two.

A number of short paragraphs, suitable for first-year pupils, was composed. In these the vocabulary was limited to the words common to the thirteen beginning books listed by Henmon.[3] Simple inflectional and syntactical material was used such as every beginner studies well before the end of his elementary work.

[3] "The Measurement of Ability in Latin," *Journal of Educational Psychology*, 8:515-38, November, 1917.

For other passages, selections from classical Latin, simplified and shortened, were used. A considerable amount of text material similar to Cæsar's *Gallic War* was read and some thirty or forty passages were tentatively selected. Several thought questions were arranged for each selection. This was an easier task for paragraphs composed for the purpose than for selections from an ancient text. The questions were so arranged as not to depend on a single sentence or on the previous context, and they were unambiguous and direct with answers which were brief and easily scored.

The preliminary form was finally reduced to four selections taken from elementary paragraphs, and eleven selections taken from Cæsar, with a short paragraph from Cicero added by way of experiment. These paragraphs were mimeographed on separate sheets. At first the questions were placed above the reading matter. We feared that if the questions followed the paragraphs, some conscientious pupils would not re-read the paragraphs while others would undoubtedly do so. Accordingly, we put the questions first in the desire to make the test as fair as possible for all participants. Careful observation and questioning, however, failed to show any advantage in this method, and it was abandoned. At the top of each paragraph in all later forms of the test, this direction was given: "Read this and then write the answer. Read it again if you need to."

A preliminary form of the test was given to the pupils at the University of Iowa High School and the high schools of Iowa City; at the high schools of Elyria and Oberlin, Ohio; and to a class at the University of Iowa which had been reading Cicero for a semester and Virgil for half a semester. The total number of pupils tested on the different paragraphs varied from 93 to 277.

From the carefully timed records kept for each section of the test, it was found that much shorter periods of time were needed than had been anticipated. From the scoring of the papers it was decided that certain paragraphs were unsuitable. These were dropped or altered.

Since these preliminary forms of the test had consisted essentially of elementary material or selections from Cæsar, we had no method of sifting the first- and second-year pupils from those of the third and fourth year. Accordingly, the writings of Cicero (other than orations found in school editions) and the works of several other prose writers were examined in search of suitable passages. The works of Virgil, Ovid, and other poets were examined for material in verse form. As a result, ten passages, six in prose and four in verse, were chosen. These selections were then given to 58 pupils in the Cedar Rapids High School, the high school at the University of Iowa, and to the pupils of a Virgil class at the University. All these pupils were in the sixth, seventh, or eighth semester of the study of Latin.

After these papers were scored and the results tabulated, the work of arranging the material into approximately final form was begun. Twenty passages out of the twenty-six previously used were combined in two forms, A and B, of the test. Each form contained in succession two elementary paragraphs, composed for the purpose, two selections from Cæsar, one from Cicero, one poetical selection, and a third selection from Cæsar, a second from Cicero, a second poetical quotation, and a fourth selection from Cæsar. The stress on Cæsar was due not to the conviction that his works are necessarily the most important, but rather because elementary pupils can, under present conditions, read Cæsar more easily than other classical Latin.

These two forms of the test were printed and were given on successive days between April 20 and 25, 1922, in four Chicago high schools: Hyde Park, Englewood, Nicholas Senn, and the University High School. The time records kept and the scores showed convincingly that the test was by no means too long. Many pupils finished before the time was up, and many made high scores. This was a surprise to us and to the half dozen well-known Latin teachers and other specialists to whom the tests were given for comment. These experts thought that the test was too long and too hard, but the actual papers of the pupils disproved this and showed that the test could be more justly

criticized as too easy for the advanced pupils rather than as too difficult for the elementary pupils. One sixth-semester pupil and three eighth-semester pupils made perfect scores. The conflict between expert opinions and experience was probably due to the fact that the experts were thinking of the passages from the standpoint of translation rather than that of comprehension. As a translation test the Ullman-Kirby Latin Comprehension Test would be rather difficult.

In scoring the preliminary forms, A and B, a key including right and wrong answers made up from the answers received in the preliminary forms, was used. Additions were made to this key from the answers recorded on the Chicago papers. A statistical record of certain doubtful answers was kept and decisions regarding these answers were made with all the facts before us. At Chicago 562 pupils took Form A and 565 took Form B; the discrepancy in numbers is due to a few absences on the one day or the other. Form A proved very satisfactory; we had put our best material into it. In revising preliminary Form A only a few slight changes were made, and these were chiefly in the wording of questions. No paragraph or question was discarded.

The arrangement of the paragraphs was changed after it became clear that the pupil had time enough to reach the last paragraph. It was interesting to note how often selections from Cicero and the poetical passages were untouched, while the questions on the last paragraph, a selection from Cæsar, were answered. The rearrangement of the paragraphs was determined by their difficulty as shown by the percentage of pupils giving correct answers to the questions. The preliminary Form A, when revised, became final Form I of the test in which the arrangement of the passages stands as follows: two elementary paragraphs, four pieces from Cæsar, one poetical selection, two quotations from Cicero, and a second poetical passage.

Form B, as was anticipated, proved less satisfactory. Not only were some of its paragraphs and questions unsatisfactory, but its mean did not correspond closely enough to that of Form A, nor was its correlation with Form A as high as was

desired. After rather extended modifications of Form B, it became Form II of the test. Forms I and II were then given on successive days, June 11 and 12, 1923, to the pupils in the Chicago high schools. In all cases, Form I was given before Form II.

The following paragraph used for practice in the tests, though easier and shorter than the paragraphs of the test itself, shows their general nature.

> Duōs fīliōs agricola habēbat, Mārcum et Sextum. Magnus erat Sextus et Mārcus erat parvus. Pater parvum fīlium magis amābat et eum ad urbem saepe mittēbat.
>
> (a) Which was more loved by his father, Sextus or Marcus?
>
> .....
>
> (b) Where did his father send Marcus?
>
> .....

Table I gives the data from which the critical evaluation of these tests has been made.

TABLE I. ULLMAN-KIRBY LATIN COMPREHENSION TEST

	Semester II		Semester IV		Semester VI		Semester VIII		Semesters Combined	
	Form I	Form II	Form I	Form II	Form I	Form II	Form I	Form II	Form I	Form II
Pupils Tested	123	123	101	101	109	109	90	90	423	423
Median	11.38	10.15	17.39	18.59	21.81	23.11	25.41	25.67	17.80	19.64
Mean	11.20	10.70	17.43	18.19	21.58	23.11	24.85	24.67	18.47	18.76
Sigma	2.48	3.96	4.05	3.70	4.79	4.00	4.67	3.99	6.59	6.98
Reliability Coefficient	.53 ± .04		.65 ± .04		.57 ± .04		.71 ± .04		.85 ± .02	

The Stability of the Mean and the Standard Deviation

The data of Table I show that the means in the two forms for any semester are not more than one and one-half score units apart and that the standard deviations of the two forms for any semester are within three-fourths of a score unit of each other. The means show that Form I is slightly easier for second- and

eighth-semester pupils than Form II, but that Form II is slightly easier for fourth- and sixth-semester pupils than Form I. The fact that Form I in every case was given before Form II gives the advantage of practice to Form II. The variations in difficulty of the two forms are small. Medians on Form I for many more cases than are here reported show for each semester a variation of less than one score point from the medians given in Table I. The progress in score points from one semester to the other is marked. The gain from the second semester to the fourth is greater than from the fourth to the sixth or from the sixth to the eighth.

Reliability of the Test

To what extent will the two forms of this test tend to give a pupil the same relative position in the two distributions of his group? In other words, to what extent will the test, if used today and tomorrow, tell the same story? If each pupil gains about the same rank in one form that he gains in the other we may conclude that the two forms measure practically the same thing and that they measure it accurately. The common method of expressing this relationship of one form of a test to another is the reliability coefficient.

The reliability of the test is shown in the coefficient of correlation between Forms I and II of Table I. These run from $0.53 \pm .04$ for second semester Latin to $0.71 \pm .04$ for eighth semester. The increase is due in part no doubt to the greater number of questions answered in the higher semesters. These correlations are not as high as are desirable, but when we bear in mind that they express the relationship for a single grade, their significance is enhanced. When the four groups are correlated together the correlation rises to $0.85 \pm .02$. Kelley has pointed out that "To secure a reliability coefficient of 0.40 from a group composed of children in a single grade is probably indicative of greater, not less, reliability than to secure a reliability coefficient of 0.90 from a group composed of children from the

second to the twelfth grades."[4] This is on the assumption that from the second to the twelfth grade the spread of ability is four times what it is in one grade. In our data the correlations emphasize the same point, though the spread of ability in our groups is not so great, being only about twice as great in the four semesters recorded as in the second semester.

With this reliability of the two forms of the test determined, how accurate is the score made by an individual pupil who makes, let us say, a score of 24? That is, if we had many forms of this test to give to a group of pupils, how far would the score of 24 made on any one form differ from the average of the scores this pupil would make on the many forms? In other words, how far does his score differ from the true score? Kelley states that "The standard error in estimating True Scores from X_2 scores is equal to $\sigma\sqrt{1-r_{12}}$"[5] in which σ is the standard deviation of the form that has been given and r_{12} is the reliability coefficient of two forms of the test. Using the eighth-semester figures in Table I, σ in Form I is 4.67 and r_{12} is .71. Then substituting for these values the standard error of estimate of a true score from a score on the single form is 3.28. The probable error of estimate is .6745 of 3.28, or 2.21 score points. That is, the chances are equal that the true score of the pupil who scores 24 on the one form will not go above 26.2 nor below 21.8. This point of the reliability of any pupil's score on a test is one of great significance.

Validity of the Test

In trying to answer the question, does this test measure what it purports to test, that is, does it test ability to comprehend Latin paragraphs, no absolute criterion with which to compare it is at hand. No one will question the validity of the Latin selections, since the judgment of the most competent Latin authorities was that these passages represented well the Latin which high-school pupils ordinarily read. Moreover, Latin para-

[4] "The Reliability of Test Scores." *Journal of Educational Research*, 3:374, May, 1921.

[5] *Op. cit.* 3:372, May, 1921.

graphs and questions in English regarding them would seem to constitute a valid test of ability to comprehend Latin.

Some indirect data are at hand that throw light on the problem of validity. First, the test was given to seniors in the University of Iowa some of whom had never studied Latin. Though urged to guess at the thought of the passages and the answers to the questions, the scores of those who had not studied Latin was almost zero. Then too, the fact that added semesters of Latin tend to bring increased scores on the tests is some evidence that it measures ability to get thought from the printed Latin page.

Perhaps the achievement that is closest to the one measured by this test is translation. Were there some accurate measure of translation, one would expect a high, though not a perfect, correlation between scores in translation and scores on a valid comprehension test. While marks have been given on translation for centuries, no objective means of scoring this type of achievement had been devised which the authors could use for validation purposes.

Some correlations with marks whose reliability is not known are the best that can be offered. For 200 second-year pupils in New York State who had taken Form I of the comprehension test and the Regents' Examinations, the correlation was $0.61 \pm .03$; for 90 third-year pupils, $0.66 \pm .04$; for 63 fourth-year pupils, $0.48 \pm .06$. Wood comments favorably on the value of Regents' marks. He found these much more highly predictive of college success than secondary-school marks. He says, "All the high schools in the country should report their grades in units as comparable as are those used by the high schools within New York State."[6] Just how reliable these Regents' marks in Latin are, the writers have no means at present of knowing. Let us suppose they have a reliability coefficient of 0.81. Then the highest correlation, except in spurious chance cases, that could be expected between them and our comprehension test in second-year Latin would be 0.84. Wood states that,

[6] Wood, B. D. *Measurement in Higher Education.* Yonkers-on-Hudson, World Book Company, 1923. p. 84.

"The maximum possible correlations (except by chance) between two variables is the product of the square roots of their reliability coefficients."[7] Anyone familiar with the Regents' examinations knows that they test many other phases of Latin than direct reading ability. Since their reliability is probably not greater than 0.81, while the reliability of our test is on the average 0.61, higher correlations than those reported above could hardly be expected.

It had been part of our plan to use teachers' marks assigned solely on ability to read and comprehend Latin, as one criterion for the validation of our test. The unreliability of such marks, however, is known to be so great as to render this criterion almost valueless. That the test measures the ability which it purports to measure can be settled only when we have a valid criterion with which to compare it. That it is a considerable improvement over previous means of testing the ability to comprehend the Latin page, we believe to be true.

[7] Wood, B. D. *Op. cit.* p. 49.

Editorials

THE NATIONAL POINT OF VIEW IN TEACHER TRAINING

In any comprehensive investigation of conditions in education, either as they are or as they ought to be, one is practically certain before long to run into embarrassment due to the existence of state lines. We say education is a state function, and indeed it is. The laws sustain this position and the courts affirm it. It is necessary, however, in education as in other enterprises, to face the facts. One fact is that no state exists unto itself. It is a state in a federation of other states. In other words, it is a state plus all the other states.

We may therefore hold as strongly as we please to the doctrine of states' rights and yet recognize that in the exercise of those rights the existence of other states cannot be ignored. Indeed, to ignore the rights, the privileges, and interests of the rest of the country is to leave out of account important data with reference to the action of the individual state.

We have lately had this matter brought to our attention in connection with an investigation of the supply of trained teachers in relation to the demand for them. The movement of teachers has no great respect for state lines. It would be contrary to the best interests of the schools if it did. Yet, as things are now managed, large expense is incurred for the training of teachers for certain types of service, and courses appertaining to these types of service are weakened because those who administer these matters are either unable or unwilling to take the national point of view.

Each year a certain state appoints fifteen or twenty new teachers of physical education, a dozen new teachers of art, and

eight or ten teachers of the deaf. These teachers must be trained, and since their service is of a peculiar nature, graduation from a general teacher-training course will not suffice. Careful, specialized courses are required in order to meet these conditions. If a course in the teaching of art or of physical education is to be really effective it must be largely independent of the usual offerings of a teacher's course. General methods, for example, will not be nearly as effective as methods in art or physical education.

Now, a state which requires but a few teachers of a particular type must either import them from another state or train them at very great expense. If it imports them, it appropriates from another state teaching material to which, on the states' rights basis, it has no title. At this point, however, the supporters of local control abandon their notion of states' rights and assert expansively that "we are all one country anyway." Quite so; the national point of view puts in an appearance, but it does so to justify a selfish policy. To be sure, a state may send as many teachers to other states as it receives; but there is normally a "balance of trade" in respect to teachers between any two states, and in virtue of it one of the two states suffers constant loss. Among all the states only a certain definite number of teachers are trained each year. It is against public policy that certain states should get more than their rightful quota of them.

If, on the other hand, a state which requires but a few teachers of a given type tries to train them in its own institutions, it will have to provide an establishment out of all proportion to the benefit it receives. If it needs ten drawing teachers a year, it will have to furnish housing, equipment, and an instructional staff for a much larger group. If these facilities are furnished, the cost of the ten teachers who annually enter the schools is excessive. If they are not, the teachers will be poorly trained.

Here the national point of view should come in. There are probably two or three adjacent states whose needs for drawing teachers would, if pooled, justify the maintenance in one of these states of a first-class department for the training of such teach-

ers. The expense might be borne jointly and the benefits shared. The same might be true of other types of teachers. When, however, we transcend the particular state and consider cooperation among several states, we find we have no machinery for putting such cooperation into effect. Only occasionally is a project carried on jointly by more than one state—unless the matter is handled by the federal government.

The *nation* therefore has a part to play in teacher training. What that part should be, we do not attempt to say. But we believe it should be something more than the benevolent bestowal of advice. When a state trains, as one did, thirty kindergartners and gets only one of them for its kindergartens, and when, as was actually the case, all the graduates of a certain department for training teachers of physical education accept positions immediately in another state, something in the way of federal supervision would seem to be necessary.

When a teacher desires, after preliminary service in the state in which she was trained, to transfer to another state, much annoyance may be experienced in obtaining the necessary teaching certificate. Some uniformity of procedure is desirable. The federal government is the only agency in a position to act effectively.

Some states enact laws which discriminate unfairly against other states. For example, not long ago a law was in effect in one state—and very likely it still is—which required that teachers who had the minimum amount of professional training should have a year of experience before they could teach in the state in question. The only way in which one might qualify under this law was to teach outside of the state for a year and then return. By what right did this state inflict upon other states the apprenticeship of its least trained teachers?

These are some of the considerations which, after a rather prolonged study of the state administration of teacher training, have led us to believe that the nation as well as the state has a part to play in this matter. This is one of the reasons why we favor a bill to create a federal department of education.

Reviews and Abstracts

PORTEUS, S. D. *Studies in Mental Deviations.* Vineland, N. J.: The Training School, 1922. 276 pp. (Publications of the Training School at Vineland, N. J., Department of Research, No. 24, October, 1922.)

This volume presents "a series of studies apparently disconnected but all bearing a close relation to mental diagnosis." It is a collection of research contributions on various aspects of feeble-mindedness most of which have previously appeared in the technical journals. Several new investigations appear here for the first time whose value, we fear, may be lost without journal publication.

Following a general statement of aims, methods, materials, and classification of subjects, the book presents the results of the author's experimental studies of normal and abnormal children and adults in the field of anthropometry (cephalometry), mental tests (Binet Scale, Porteus maze, Porteus form and assembling tests), personality study (social rating scale), achievement (educational rating scale, industrial rating scale), and case studies. Particular attention is paid to sex differences and to differences associated with mental abnormalities. Each separate study is a contribution to technique and norms. The unifying bond in such apparently diverse investigations is the commendable attempt to improve upon current methods of the diagnosis of mental deviations by extending the fields and methods of clinical psychological examination, the application of which are not limited to mental deficiency alone.

A critical evaluation of the author's several studies is of perhaps little interest to the general reader; the technical student will no doubt prefer to draw his own conclusions. Suffice it for present purposes that the several presentations are clear and direct, although by no means always convincing or above debate. Thus the author's psychology is better than his statistics and one regrets that in order to make a case for his argument he is so frequently uncharitable to others.

Considering the volume as a whole, one observes a lack of integration and of generalization desirable to avoid the criticism that the book is a bound volume of reprints. The book is neither a manual of clinical procedure nor an exposition of clinical theory and technique, although something of both. Yet one is impressed with the distinct contribution to both technique and theory in the separate chapters. There also is merit in the continued insistence on the utilization of all possible data, in the analytical approach, in the experimental attitude, and in the absence of complacency. The insistence on studying the whole *human* individual is particularly gratifying, and the final chapter on "Children—not Cases" is a literary and scientific classic.

The reader may confidently expect to find a positive contribution to the scientific study of feeble-mindedness and a constructive addition to technique in mental diagnosis. Obviously the applications of the author's results are not limited to the field of abnormal psychology, but are specially important in the clinical study of all individuals in their educational and social aspects.

EDGAR A. DOLL

Ohio State University

DAVIS, CALVIN O. *Junior High School Education.* Yonkers-on-Hudson, World Book Company, 1924. 451 pp.

Earlier books on the junior high school have devoted themselves in the main to particular phases of junior high-school work, but under the title, *Junior High School Education*, Prof. Calvin O. Davis of The University of Michigan, has collected what is in effect a compendium of junior high-school information.

This new book follows Briggs in giving, on practically every topic, the views and practices of a great many junior high-school principals and superintendents the country over; Koos in attention to the administrative side of junior high-school work; Van Denburg in giving one chapter to each of the major lines of junior high-school study, beginning with physical training and health and ending with morals and manners. It resembles Johnston's book in its chapter of some twenty pages on "Adolescent Characteristics." The bulk of the book is made up of quotations from almost every worthy source: books, reports, bulletins, pamphlets, articles, addresses written by mature students and often by committees interested in the junior high-school movement. There is indeed hardly any phase of junior high-school theory or practice—with one conspicuous exception noted below—that does not receive attention.

For the most part the author's own contribution consists in his discriminating selection of the views of others. The author's interpretation of these views would have made the book more interesting, but he is for the most part content with quotations which show the present trend.

This book is no more easy reading than an encyclopedia is, but every profound student of the junior high-school movement should be acquainted with the information here gathered under one cover.

We regret, however, to note the almost entire absence of any mention of the use that is being made of the newer group tests of so-called "general intelligence." The use of these tests in forming homogeneous mental groups, in forecasting the probabilities of the school success of junior high-school pupils, is not even mentioned. In this respect the date of publication might as well have been 1920 as 1924.

There is, however, no doubt that this book, though incomplete, is still the most comprehensive one yet published in the junior high school.

SARA A. NAGLE

Board of Education, New York City

The Junior High Schools of Rochester. New York, Board of Education, 1923. 230 pp.

This volume is a report prepared by the Junior High School Council of the City of Rochester and attempts in a highly effective and interesting manner to tell the main facts concerning the development of the junior high-school movement in that city.

Superintendent Weet, in a brief introduction, states frankly that the junior high school is costing the city considerably more than did the older type of organization on the eight-four plan, and declares that the primary purpose of the report is to justify these increased expenditures before the public.

The report is divided into five main parts: The Origin and Growth of the Junior High-School Plan; Organization, Administration, and Supervision; Curricula and Courses of Study; Results; and Costs.

Rochester has organized its junior high-school work "primarily for the purposes of general education as opposed to training for a specific trade or occupation." Neither has it provided any considerable vocational education for pupils of the senior high schools. Nevertheless, Superintendent Weet, and the Council seem to be convinced that equal educational opportunity does not mean the same opportunity for all pupils, that the printed page is only *one* pathway to an education for boys and girls, and that for those individuals who by interest, ability, and circumstances are likely to find their life work in industry and not in professional careers much more quasi-vocational work should be provided than has been provided in the past—either in the junior high school, senior high school, or special vocational schools.

To organize and administer the work of the junior high schools so that full justice to pupils and to society may be conserved has been one of the outstanding tasks of the school men and women of Rochester, and the efforts of these persons to accomplish this task constitute the content and value of the report.

Among the policies Rochester seems to have found most helpful in reaching its objectives are the following:

1. Characteristic curricula or courses of study adapted to the needs and abilities of varying groups of pupils were formulated by teachers' committees.
2. Many experienced upper-grade teachers of the older grammar school are employed in the junior high school, but these teachers are required to prepare themselves for the new position.
3. The principals of the junior high schools are former successful principals of elementary schools.
4. They have separate departmental heads in the junior high school but utilize the experience of senior high-school departmental heads in connection with the organization, supervision, and administration of departmental work in the junior high school.
5. Close attention is given to the intellectual, social, and physical needs and practices of pupils from the moment they enter the junior high school till the time they leave.

This report is filled with concrete material of many sorts. One large division of the report describes the duties of the various school officers and

committees, tells the numerous types of special school activities; discusses the plan of supervised study employed, defines the socialized recitation which is set as an ideal of class procedure, and gives detailed outline of the various courses of study and curricula offered.

In computing the educational results of the junior high schools, three sets of materials were used. These were:

1. The semester reports gathered for three years and covering all ninth-grade subjects.
2. Comparative tests in Latin and mathematics.
3. The complete senior high-school records of all junior high-school graduates entering a given senior high school.

On these three bases the conclusions of the Council are: (1) The attainments of junior high-school pupils are high. (2) More than 50 percent of these pupils who enter the senior high schools sustain themselves with marked credit. (3) There is a striking uniformity of high marks in English, Latin, and mathematics in the two types of school.

A similar attempt is made to judge of the social results, the vocational results, and the holding power of the junior high school. Naturally here the measurements are not as accurate nor the deductions as positive. Nevertheless the following conclusions seem fairly to be reached:

1. That junior high-school pupils are being given health training that is likely to function in later life vastly more effectively than anything that was formerly done in the old-type school organization.
2. That they are being provided with socializing experiences that seem bound to "insure the exercise of those habits and attitudes which have permanent value in social and civic life."
3. That those boys and girls who have but a limited time to spend in school and who seek to secure elementary specific training for commercial and industrial positions are being prepared for such positions in a very adequate way.
4. That many more pupils of all types are being held in school for longer periods than was the case under the older eight-four type of school organization.

Taken as a whole, the report is an exceptionally interesting and valuable piece of work. It is essentially scientific in its plan and procedure. The descriptions of the building plans, the room arrangement, the kinds and costs of equipment, and other pertinent aspects of the junior high-school problems are definite and concrete. No school administrator who is contemplating the organization of a junior high school can afford not to be acquainted with this report. Indeed, if every teacher in our schools of education, senior high schools, junior high schools, and elementary schools were to review the book there would be produced a clearer notion as to what the junior high-school movement really is and how it concerns each one in the entire educational system.

CALVIN O. DAVIS

University of Michigan

O'SHEA, M. V., editor. *The Child: His Nature and His Needs.* Valparaiso, Indiana, The Children's Foundation, 1924. 500 pages.

The promise made in the title is amply fulfilled by the subject matter of the book. In a single volume it sums up and applies the general principles involved in most that is known regarding the nature and the physical, intellectual, social, and moral needs of childhood and youth. It aims to bridge the gulf between the work of the investigator and that of the practitioner.

Prepared under the editorship of M. V. O'Shea, it consists of twenty-one chapters written by sixteen well-known investigators and practitioners in the field of applied psychology. It includes such pertinent topics as health and physical development, intellectual superiority and inferiority, mental hygiene, sense defects, social traits, problems of adolescence, and the changing objectives and methods in the education of children. The book is conveniently indexed, interestingly illustrated, and is provided with biographical data and copious bibliographies carefully classified.

Discussion of the material would involve discussion of the investigations of sixteen authorities whose work is already too widely known for casual critical comment. It is not necessary to agree with the details of every statement or to reconcile any slight differences of opinion to recognize the invaluable service of such a rich mine of current information.

It embodies the rare combination of being scientifically sound and, at the same time, sufficiently comprehensible, readable, and entertaining to hold the interest of any reader who is interested in children. Because of its condensed character and infinite variety, it demands more than one perusal for satisfactory assimilation.

As the reader doubtless knows, The Childrens Foundation was granted a charter by the State of Indiana in 1921, having as its object the study of children and the dissemination of knowledge promotive of their well-being. The Foundation is a corporation "not for profit." The trustees, therefore, have provided funds for the gratuitous distribution of the book to a number of institutions and individuals who are rendering distinguished service to the cause of childhood. To others it is made available by a nominal contribution to the Publication Fund of the Foundation.

The book is a tremendous undertaking successfully accomplished. It arouses the reader's anticipation of the second contribution of the Foundation, which is promised for the spring of 1925.

GRACE E. BIRD

Rhode Island College of Education

DRESSER, HORATIO W. *Psychology in Theory and Application.* New York, Thomas Y. Crowell Company, 1924. 713 pp.

The purpose of the book, according to its preface, is "to coordinate the various divisions and applications of psychology so far as they belong under the head of science." This is obviously a large order but the author attempts to cover most of the fields of general, abnormal, business, and social psychology.

Part I presents general psychology in rather conventional fashion with many quotations from James, Warren, Woodworth, McDougall, and others, especially McDougall. The emphasis is purposive rather than behavioristic.

Part II discusses the psychology of the hidden self. There is a jumbled chapter on psychotherapy but a good summary of Freud's theories and a little critique thereof.

Part III deals with vocational and industrial psychology, namely, mental tests, mental efficiency, creative impulse, business success, advertising, and selling. The discussion of mental tests is confined entirely to intelligence tests with no mention of the procedure of correlating tests with some vocational criterion. One chapter is inspirational rather than psychological. There is practically no mention of the experimental work on advertising. Illustrative examples would improve this and the preceding part.

In Part IV social psychology is approached through the instincts—gregariousness followed by McDougall's list. The role of mental defect and mental conflict in delinquency and crime is described; and the psychology of testimony, propagandism, and war is discussed. "The real lovers of peace in the world are those who are trying to find a complete moral equivalent for conflict." While this part of the book occasionally forsakes psychology for such things as "the better self" or "loyalty to loyalty" it is probably the best of the five.

Part V is devoted to social organization. This includes social conflicts, domestic life, play, education, and religion. The discussion of educational psychology is very brief, about ten pages compared with over one hundred pages devoted each to abnormal and business psychology. The final chapter raises certain critical problems such as agnosticism, pragmatism, and the locus of values to which the author claims psychology ought to contribute. No such contribution is, however, likely "until psychologists overcome the habit of keeping so close to the ground, devoting nearly all the time to studying the beginnings of mental life and leaving scarcely an hour for our higher nature." The reviewer is inclined to think that the ground is the only safe place in the present status of the science, and that, if the book's mission is to induce psychologists of the present generation to forsake the world of facts for the world of values, it will fail.

The book is well written, systematic for the most part, and replete with references. The author has evidently read extensively, not merely books but recent periodical literature. He quotes judiciously. Obviously such a large undertaking must omit much and confine itself to a selection of important points. This selection seems to have been more wisely made in the topics where there is no available experimental material. In the others one misses the laboratory attitude and the statistical approach.

The book will probably be of interest to the layman who does not have time to read intensively in any phase of psychology but merely wishes a general picture. Students who have had a few introductory courses in psychology may find it profitable to read portions of the book by way of review before proceeding to specialized work. Advanced students may find the references use-

ful and some of the later chapters might furnish a basis for discussions in social psychology, ethics, and sociology.

HAROLD E. BURTT

Ohio State University

COOPER, HOMER E. *Cost of Training Teachers.* Baltimore, Warwich and York, 1924. 111 pp.

This book has direct significance for those interested and engaged in the administrative work of training teachers. The author has attempted to ascertain the cost of teacher training for the state of New York with the view of giving information for the establishing of a policy for the training of its teachers. He has tabulated the total cost of maintaining teacher-training institutions in the state. An attempt is made to determine the turnover of the teaching body by the number of certificates issued each year. The author endeavors to estimate the average length of service by the use of the following formula:

$$\frac{\text{No of certificates used}}{\text{No. of certificates issued in a year}} = \text{Average years of service}$$

He finds by this means that the average teaching service for the teachers of New York state to be 6.75 years. This can only be an assumption and may approximate the facts, but his findings do not agree with any of the other writers on the subject. He states that the war period did not effect the figures.

The book is well illustrated with tables and graphs. It contains many assumptions which on first reading seem doubtful, but further research may establish them as facts. There is an interesting appendix and a selected bibliography.

C. E. BENSON

School of Education, New York University

CALDWELL, OTIS W., AND COURTIS, S. A. *Then and Now in Education.* Yonkers-on-Hudson, World Book Company, 1924. 400 pp.

In 1845 the Boston School Committee decided to give a written examination to the upper class of elementary-school pupils—those corresponding to our present eighth grade—in history, astronomy, arithmetic, and geography instead of the time-honored oral quiz. The questions asked in these examinations and the answers made by the pupils are the data used by the authors for the "Then" in education. The examinations were given to a selected group of the upper class, "the flower of the Boston schools." The masters would not permit the examination of all the upper class. The percents of success made by the brightest pupils were as follows: history, 45 percent; astronomy, 39 percent; arithmetic, 35 percent; geography, 34 percent; and the average on all subjects 30 percent.

The world in which the children of 1845 lived had few points in common with the world of the eighth-grade pupil of 1919. Hence the old examinations were adjusted to meet modern conditions by selecting five questions in each subject "which seemed to have possibilities for twentieth-century children."

The answers to the questions given to unselected pupils in the eighth grades of a number of cities constitute the "Now" in education.

In comparing the results of "Then" and "Now" the point is made that the Boston schools of 1845 retained their bright pupils for a longer time than is the case in the elementary school of today. It appears, then, that the comparison is made between the results of teaching in 1845 as affecting the brightest pupils, and the results of teaching in 1919 as affecting the unselected lower 40 percent of approximately the same age group.

The outstanding conclusions drawn from the repetition of the tests are that: first, the present-day children made lower scores on pure memory and abstract skill questions and higher scores on thought and meaningful questions. "On the whole, therefore, the 1919 children must be judged to be more intelligent, more able to think, and to apply knowledge than the 1845 children, but they have less control over abstract skill and concepts." Second, the efficiency of present instruction, at best, although higher than in 1845 is still far from satisfactory.

Following these conclusions is an extended discussion of the general administrative conditions under which the schools of "Then" and "Now" did their work. Particular attention is given to the school system of Detroit as a basis for comparison and prophecy. The latter half of the book gives the source material upon which the analysis of the results of the 1845 survey is based.

This book is interesting and well written. It affords a really scientific answer to the criticisms so frequently made that schools of today are not doing as good work as the schools did in the good old days. If the critic's conception of the purpose of education is that of 1845, of course the schools are not achieving what they should; on the other hand, if what we want is initiative, resourcefulness, thinking to a purpose, the critic will have some difficulty in disposing of the conclusions reached by a comparison of the educational results of 1845 and 1919.

GEORGE L. JACKSON

University of Michigan

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

Dr. George Willard Frasier will be formally inducted into office as president of Colorado State Teachers College on Friday, November 21, at Greeley, Colorado.

The inaugural ceremony will take place at 10:30 in the morning, following an academic procession in which leading educators from different parts of the country, prominent citizens of Colorado, and students will take part. The governor of the state or his representative will preside at the inauguration and Dr. Charles McKenny, president of Michigan State Normal College, will deliver the inaugural address. The Very Rev. Dean B. W. Bonell of St. John's Theological College, will deliver the invocation.

Ethan Allen Cross, dean of the college, will deliver the address of welcome and President Frasier will respond. H. V. Kepner, president of the Board of Trustees, will present the new president with the keys to the college.

One thousand invitations are being issued and it is believed that the occasion will bring to the campus of Colorado State Teachers College the largest and most important gathering of educators ever assembled in the west.

A banquet will follow the ceremonies, at which Dr. Guy Maxwell, president of Winona State Teachers College, Winona, Michigan, will act as toastmaster. It will be at the banquet that general addresses will be made by distinguished visitors. In the evening a public reception will be held.

Form B of the Chapman-Cook Speed of Reading Test for grades IV to VIII appeared during the summer. Those who are familiar with the Form A will doubtless welcome an alternative form for use in measuring progress. The publishers are the J. B. Lippincott Company.

Doctor Frederick L. Whitney, for some time director of the Training School at the State Teachers College at Moorhead, Minnesota, became the director of research at the Colorado State Teachers College, Greeley, with the beginning of the current school year.

Children's Book Week has been announced for November 9 to 15, inclusive. Posters and pamphlets can be secured from the local book stores and the public library, or from the National Association of Book Publishers, 334 Fifth Avenue, New York City. More than 3,500 cities and villages observed Book Week in 1923.

Doctor R. D. Porteus, director of research, Training School, Vineland, New Jersey, has issued a guide to his maze test. The guide contains a general introduction, general instructions, interpretation of test results, and directions for application and scoring.

The World Book Company has published a "Universal Percentile Graph" by Arthur S. Otis. The graph provides an easy means of recording scores in such a manner that medians, quartiles, percentile ranks, as well as range of distribution are easily seen. It is undoubtedly a time saver for those who have a large number of classes of considerable size for which these varying items are decided on test results.

"If culture does happen to be found in a college today it is because the institution harbors a professor who is a 'Pedagogical bootlegger.'" Thus Doctor C. G. Shaw, professor of philosophy at New York University, is quoted in the *New York Sun*. From the sundry remarks attributed to Doctor Shaw it is evident that those courses given in colleges which contribute directly to vocational work and home life, regardless of how much they make for happiness and appreciation, are not cultural.

During the summer the World Book Company announced two new tests available for use in the coming year, the Ruch-Cossmann Biology Test for High Schools and Colleges and the White Latin Test for High-School Students. The biology test involves multiple-choice, best reason, matching of name and drawing, and sentence completion. Two forms were available and tentative norms have been established. The White Latin Test consists of two parts, the Latin vocabulary with multiple-choice of English words, and sentence translation in which a number of translations were given and the student selects the correct one. Two forms and tentative norms are available for this test also.

The general conclusion derived from studies to determine what mathematics a person needs to meet life situations outside of school, is that one needs to know:

1. How to add, subtract, multiply, divide whole numbers, common and decimal fractions.
2. Percentage with applications to saving and loaning money, mortgages, modern banking methods, building and loan associations, keeping simple accounts, investing money, bonds as investments, real estate as investments, mark of a good investment, taxes, levies, public expenditures, profits in different lines of business and life insurance as protection and investment.

3. Practical measurements, such as liquids, solids, surfaces, and time.
4. Graphical representation of data including the use of frequency distributions and central tendencies.
5. Simple equations.

Mr. William H. Cunningham of the High School of Commerce of Boston has written a pamphlet on good manners and conduct, which has been issued by the Board of Superintendents of that city. It is written for the student, and in addition to the general obligations which a student owes to parents, teachers, fellow-students, and self, it contains a large number of specific statements concerning proper conduct in numerous situations which come into student life. If one assumes that proper conduct in social situations must be learned by children, then it would seem necessarily to follow that some provision must be made for learning opportunity. Whether or not Mr. Cunningham's method of presenting the socially acceptable action to the student is the best way to guarantee their education, may be a debatable question; but that in some manner our adolescent youth should learn what is considered socially acceptable, is self-evident.

Dr. J. Carleton Bell, of the Maxwell Training School for Teachers, Brooklyn, New York, has been appointed assistant professor of education in the College of the City of New York. Dr. Bell was the founder and for many years the managing editor of the *Journal of Educational Psychology*, and was the organizer and is the secretary-treasurer of the New York Society for the Experimental Study of Education. In his new position he will be concerned particularly with the training of teachers for high schools, and much of his attention will be devoted to the development of measurements in high-school subjects. He hopes to establish in the School of Education at the College of the City of New York a bureau of high-school service, which will function in connection with the city high schools and assist them in the solution of their problems.

President Ryan of the New York City Board of Education announces that "on July 23 a cooperative, constructive school survey was started." Fifty thousand dollars was voted last year for this purpose by the Board of Estimate and Apportionment. Superintendent E. C. Hartwell of Buffalo, Superintendent W. H. Holmes of Mt. Vernon, Professor R. H. Jordan of Cornell, Mr. A. L. Weeks, a specialist in school-building efficiency, and Dean John W. Withers of New York University are among the experts who will work under the general direction of W. H. Allen of the Institute of Public Service. President Ryan further states that:

> We have discovered it is impossible for unpaid, part-time, commissioners, each busy with his own business, to be sure we are doing our best with the taxpayers' money and the best for the city's children unless we work out methods of securing needed information in advance of important decisions and unless we also secure impartial reports on actions taken.

The Character Education Institution, Chevy Chase, Washington, D. C., has sent out a "Five-Point Plan for Character Education in Elementary-School Classrooms." The plan includes (1) a classroom club to be called Uncle Sam's Boys and Girls; (2) a children's morality code, which is intended to assist teachers and children in reaching a clear understanding of laws of conduct; (3) character diagnosis, which will assist children in a true understanding of their personal strong and weak points; (4) character projects, which is designed to give exercise in right conduct; and (5) a report to parents on the character rating of children. In this report teachers are expected to check in terms of "always," "nearly always," "half-and-half," and "seldom," the degree to which children are living up to the code of conduct on the following items: honesty and truthfulness, cooperation, kindness, initiative and ambition, public spirit and patriotism, justice, self-control and poise, interest, industry and energy, concentration, executive ability and purpose, muscular control, and progress.

The Eyesight Conservation Council of America has issued a preliminary report on the eyesight of school children in the public schools of the United States. On the basis of an extensive survey covering the results of the testing of over ten million school children during the past sixteen years and of about four million during 1923, the Council reaches the conclusion that less than 17 percent of the total number of school children in the United States were given eyesight tests during 1923 and that approximately 25 percent of all school children have defective eyesight, either corrected or uncorrected. The survey further shows that the eye tests in a large number of our schools are very superficial by the fact that certain large cities report less than 10 percent having defective vision either corrected or uncorrected; while others report as high as 40, 50, and 58 percent. The report also points out that laws are not required to establish state-wide programs of eyesight conservation in the schools. Only 20 states have statutes providing for eye tests and a systematic effort is being made to enforce the laws in only 8 states.

The Committee on Comfort, Acuity, and Speed of Reading of the Council has set up a large number of research problems dealing with printed material, typewritten material, reading, and special related subjects. Such problems as size of book, length of line, width of margins, size of type, spacing between letters, words, and lines, colored printing, colored paper, brightness, glare, etc. need careful scientific solution in order that eyesight may be properly conserved.

The Downey Will-Temperament Test in the Normal School

The Downey Group Will-Temperament Test was given to 47 sophomores in the writer's classes at the Bellingham State Normal School. The scores made on the individual items of the Downey test were compared with the grades earned by the students in practice teaching, education and psychology, art, and penmanship. "Practice teaching" was chosen because the work of the normal school leads to this subject, which in turn should indicate the value of the school's product; "education," consisting of professional subjects, for they

distinguish the aim of the normal school from that of an academic institution; "art," that one might have a comparison with work distinct from the academic and educational; "penmanship," since writing enters so prominently in the Downey test.

The number of cases is not the same for each of the four series of correlations for all of the group had not had penmanship, art, or practice teaching. The letter grades given in scholarship reports were transmuted into numbers and the averages on the basis of credit hours were used in the calculation of the coefficients of correlation.

TABLE I. COEFFICIENTS OF CORRELATIONS: SCORES ON ITEMS OF DOWNEY WILL-TEMPERAMENT TEST AND NORMAL-SCHOOL GRADES

Trait	Practice Teaching	Education	Art	Penmanship
A	+.010 ± .123	−.287 ± .089	−.023 ± .123	+.180 ± .109
B	+.242 ± .116	−.040 ± .097	+.179 ± .120	+.275 ± .105
C	−.201 ± .118	−.228 ± .093	+.145 ± .121	−.013 ± .114
D	−.116 ± .121	−.19 ± .094	+.055 ± .122	−.010 ± .114
E	−.177 ± .118	+.044 ± .097	−.403 ± .100	−.175 ± .110
F	−.175 ± .118	+.019 ± .098	−.343 ± .108	+.055 ± .113
G	−.007 ± .123	+.285 ± .087	−.010 ± .123	−.139 ± .111
H	−.134 ± .121	+.0017 ± .098	+.178 ± .119	+.046 ± .113
I	+.135 ± .121	−.144 ± .096	−.058 ± .122	−.232 ± .107
J	+.274 ± .112	+.061 ± .097	+.477 ± .095	+.226 ± .108
K	−.198 ± .118	−.065 ± .097	−.051 ± .122	−.095 ± .112
L	+.071 ± .121	−.158 ± .095	−.212 ± .118	+.011 ± .114
N	30	47	30	35

Note: The traits in the test as named by Dr. Downey are as follows:

A. Speed of Movement
B. Freedom from Load
C. Flexibility
D. Speed of Decision
E. Motor Impulsion
F. Self-confidence
G. Non-compliance
H. Finality of Judgment
I. Motor Inhibition
J. Interest in Detail
K. Coordination of Impulses
L. Volitional Perseveration

For a complete description of these see Dr. Downey's *The Will-Temperament and Its Testing* (Yonkers-on-Hudson, World Book Co., 1923).

Table I is to be read as follows: the coefficient of correlation between scores on Trait A of the Downey test and grades earned in practice teaching is +.010 ±.123; between scores on Trait A and grades received in education, −.287 ±.089, and so on. In most cases the coefficients of correlation are very low with such relatively high probable errors that the majority of the coefficients cannot be said to be satisfactory.

If the tests measure what they purport to measure, the lack of contribution made by these traits to success in practice teaching as defined by normal-school supervisors is an interesting conjecture. That "freedom from load" and "interest in detail" aid in securing good grades in practice teaching might well agree with a priori judgment; but that "flexibility" and "coordination of impulses" are a hindrance would seem contradictory. Neither

slowness nor quickness of movement nor speed of decision seem to have an advantage.

The students who do not comply with instructions because they think that they themselves are right ("non-compliance") would seem to have a slight advantage in getting grades in education and psychology over those who do comply. The students characterized by flexibility and quickness of movement and speed of decision work under a slight handicap.

In art work the most striking correlation is with "interest in detail," agreeing well with subjective judgment. Success in art classes involves a great deal of attention to the details of hand work and "busy work" of the type found in the elementary school. That "motor impulsion: impetuosity and energy of reaction" should mean lower grades might well be probable; but should "self-confidence" be a handicap in securing good grades in the Art Department?

The highest correlations with penmanship grades—though low—would seem to be logical, namely, that "freedom from load" and "interest in detail" help, while "motor inhibition" detracts from earning high grades.

In general the traits as measured by the Downey Group Will-Temperament Test offer little indication of what might contribute towards success in practice teaching, education, art, and penmanship in the normal school. The majority of correlation coefficients are so low that they fail to indicate any relationship. In but two cases can we say that there is a definite correlation, that is, the correlation of "interest in detail" and "motor impulsion" with art.

A comparison of the "profiles" of the students of the group who had had practice teaching failed to reveal any clear distinction upon which to base any prediction as to success as student teachers. The mean-score-profiles for those who made an "A" in all their practice teaching and those who made an average grade of "C"—passing grades are A, B, C, and D—show that the "A" group made better scores on only three traits: motor inhibition (I), interest in detail (J), volitional perseveration (L). A comparison of one of the lowest profiles in the "A" group with one of the highest in the "C" group shows the "A" individual considerably below in several items and above in only one—volitional perseveration. Individual profiles, however, show a great deal of overlapping.

On the basis of the grouping of traits as made by Dr. Downey, the mean scores would indicate that those who received "A" in their practice teaching were more "controlled, deliberate, careful," but we must remember that a number of the "C" group made just as high scores on the items I, J, K, L, the traits indicating the above characterization.

ARTHUR KOLSTEAD

Bellingham State Normal School